

THE ATLANTIC-WIDE RESEARCH PROGRAMME FOR BFT (GBYP  
Phase 12)

SHORT-TERM CONTRACT (ICCAT GBYP 12/2022)

Design- and model-based inference to estimate density, abundance  
and distribution of BFT in the Mediterranean Sea.

**Final Report**

July 2023

**CREEM, University of St Andrews**

C. G. M. Paxton, C. Oedekoven, L. Burt, M. Chudzinska

**Instituto Español de Oceanografía**

M. Pilar Tugores Ferrà, D. Alvarez-Berastegui



This project is co-funded  
by the European Union

Please cite this report as: Oedekoven *et al.* (2023). Design- and model-based inference to estimate density, abundance and distribution of BFT in the Mediterranean Sea. Report number CREEM-2023-04. Provided to ICCAT, July 2023 (Unpublished).

### Document control

Please consider this document as uncontrolled copy when printed.

Version	Date	Reason for issue	Prepared by	Checked by
1	14 July	First draft	CO, CP, LB, MC, PT, DA	FA
2	21 July	Second draft	CO, CP, MC, PT, DA	



## Contents

Introduction .....	6
Aim of the study.....	6
Roles and responsibilities of the teams involved. ....	6
Environmental variables characterizing distribution and occurrence of BFT. ....	7
Bathymetry and slope.....	8
Salinity.....	8
Chlorophyll-a.....	9
Sea surface temperature .....	9
Sea surface temperature: temporal gradient.....	9
Fronts and mesoscale activity.....	9
Methods.....	10
Overview .....	10
Aerial surveys data (2017-2022).....	10
Extracting environmental covariates.....	11
Data sources.....	11
Spatial resolutions.....	11
Primary and derived environmental variables .....	13
Data processing and derived variables .....	13
Statistical methods - Distance sampling (DS) .....	15
Survey design .....	15
Fitting the detection function.....	17
Goodness of fit.....	18
Estimating group abundance .....	19
Statistical methods - model-based inference.....	19
Probability of detection .....	19
Modelling the number of groups.....	19
Model specification.....	20
Potential issues .....	21
Model selection .....	21
Group size modelling .....	21
Estimation of the total number of individuals.....	22

Estimation of uncertainty .....	22
Results .....	22
Distance sampling (DS) .....	22
Best fitting detection functions .....	22
Goodness of fit.....	23
Summary of effort and estimates of encounter rate. ....	25
Exploration of candidate explanatory variables .....	31
Group 1: bathy, catslope, Julian .....	31
Group 2: Slope with varying spatial resolution.....	32
Group3: <i>Chlorophyll – daily product</i> .....	33
Group 4: Sea-surface temperature – daily product.....	34
Group 5: Sea-surface temperature – 7-day temporal gradient.....	35
Group 6: Sea-surface temperature – 15-day temporal gradient.....	36
Group 7: Mixed layer depth – daily product.....	37
Group 8: Salinity – daily product .....	38
Group 9: Finite size Lyapunov exponents – daily product.....	38
Group 10: Residual temperature – daily product.....	39
Group 11: Spatial gradient for sea-surface temperature and salinity – daily .....	40
Model-based analysis .....	41
Group Number Model.....	42
Best model for group size .....	47
Estimating fish numbers and density.....	47
Estimating and modelling biomass .....	65
Discussion.....	65
Design- and model-based approaches .....	66
Environmental drivers of BFT abundance.....	68
Conclusions .....	69
Acknowledgements.....	72
References .....	72
Appendix .....	74



## Introduction

### Aim of the study

The present work aims at analysing the aerial survey data collected in order to estimate biomass and abundance of Bluefin tuna (BFT) in the Mediterranean Sea in the three surveyed blocks: A, C and E to understand which environmental covariates drive seasonal and regional changes in BFT abundance and biomass. To do this, two approaches were undertaken: 1) Design-based analysis (sensu Buckland et al. 2001) of visual aerial surveys to continue time series of abundance/biomass estimates of BFT in the Mediterranean based on new surveys from 2022 and 2) Model-based analysis (sensu Hedley and Buckland 2004) of visual aerial surveys to estimate abundance and biomass of BFT in the Mediterranean in 2017-2022.

### Roles and responsibilities of the teams involved.

The teams involved in this work had the following roles in the completion of the contract goals.

CREEM was in charge of formatting of the data for both the design-based approach, summarising survey data (e.g., effort, number and location of sighted groups, group sizes), preparing the data for model-based analysis (e.g. segments) and prediction grid, sending segments and prediction grid to CN-IEO-CSIC, undertaking design-based analysis, undertaking model-based analysis including covariates provided by CN-IEO-CSIC and writing the report (in collaboration with CN-IEO-CSIC).

CN-IEO-CSIC was mainly involved in identifying, selecting, retrieving and providing environmental data for the model-based analysis as well as previous knowledge regarding Bluefin tuna spawners. Based on previous knowledge of environmental relationships between distribution and abundance of tuna spawners, CN-IEO-CSIC have selected the primary environmental variables to be used, decided the use of the same environmental dataset for all survey blocks or different datasets, decided the time and spatial resolutions, identified environmental data sources. CN-IEO-CSIC have also created repositories of primary environmental data and post-processed environmental variables. CN-IEO-CSIC extracted environmental data at sampling segments, provided and evaluation of the significance of each variable and the potential correlations, refining the indicators and providing assessment in how to avoid possible correlation among variables. CN-IEO-CSIC selected the final variables and provided them to CREEM at the prediction grid points on selected prediction spatial scale.

Below a short description of the authors of sections of the present report is included:

- Introduction
  - o Aim of the study (CREEM, IEO-CSIC)

- Environmental variables characterizing distribution and occurrence of BFT (IEO-CSIC)
- Methods
  - Overview (IEO-CSIC, CREEM)
  - Aerial surveys data (2017-2022) (IEO-CSIC, CREEM)
  - Extracting environmental covariates (IEO-CSIC)
  - Statistical methods – Distance sampling (DS) (CREEM)
  - Statistical methods – model-based inference (CREEM)
  - Estimation of the total number of individuals (CREEM)
  - Estimation of uncertainty (CREEM)
- Results (CREEM)
- Discussion
  - Design- and model-based approaches (CREEM)
  - Environmental drivers of BFT abundance (IEO-CSIC)
- Conclusions (IEO-CSIC)

### **Environmental variables characterizing distribution and occurrence of BFT.**

The selection of potential explanatory environmental variables was based on established connections between the ecology of BFT and mesoscale oceanography. All environmental variables selected have been already explored in the scientific literature from studies conducted in the three study blocks (referred to as *Blocks* in the text) (Figure 1). We only included environmental variables that are known to drive the spatial and temporal distribution of the BFT, taking care that environmental variables associated to oceanographic processes affecting stock productivity were excluded to avoid confounding effects with the interannual variations in abundance. Specifically, oceanographic processes that influence the location of BFT spawning areas included sea surface temperature, salinity, geostrophic velocities, and chlorophyll-a (Alemany et al. 2010, Reglero et al. 2012, Alvarez-Berastegui et al. 2016), and other derived variables from those essential ocean variables. Another important parameter that has been considered along the study is the spatial resolution of each of the environmental variable processed, as the scale of observation directly affects the capability to identify environmental-ecological relationships (Alvarez-Berastegui et al. 2014).

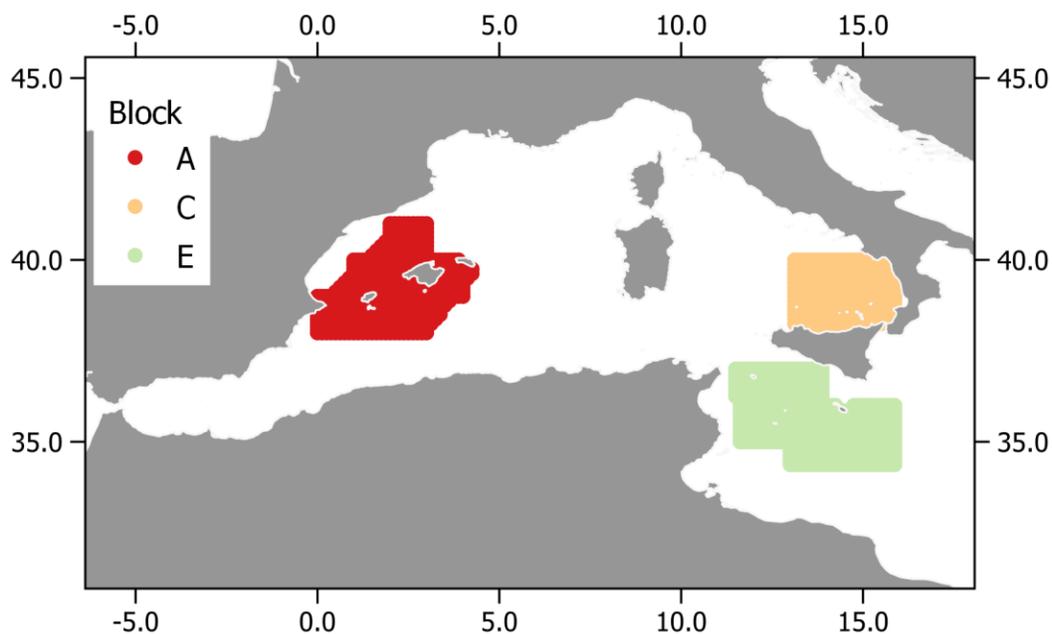


Figure 1. Depiction of Mediterranean Sea and the three study blocks (referred to as *Blocks* in the text).

### Bathymetry and slope

Different studies have related BFT spawning areas to bottom topography. In the western Mediterranean, BFT prefers spawning in deeper waters, with bottom depths between 1,250 and 2,250 m depth (Alemany et al. 2010). In the Gulf of Mexico, another prominent spawning area for BFT, spawning may occur in waters with wider ranges of bottom depths, including the shelf break, the slope and the deeper waters (Hazen et al. 2016). Indeed, hotspots for spawning BFT have been found in slope waters of the Gulf of Mexico, when electronic tagging and pelagic longline data are combined (Block et al. 2005).

Bathymetry and slope are environmental variables reflecting the topography of the bottom of the sea. BFT, spawns at the sea surface So, the potential relationship of these variables with BFT spawning are explored for potential association with other indirect effects, as the steep slopes around the Balearic archipelago define main surface currents that are associated to the spawning aggregations.

### Salinity

Salinity is associated with frontal areas and is one of the most relevant environmental variables that explains the distribution of BFT larvae in various spawning regions. Studies in the Balearic Sea (Alemany et al. 2010, Reglero et al. 2012), the Gulf of Tunisia (Koched et al. 2013) and the Gulf of Mexico (Muhling et al., 2010, 2013) have all linked salinity to the spatial distribution of BFT larvae. In the Balearic Islands, offshore mixed waters close to frontal areas at the confluence of Atlantic and Mediterranean waters, appear to be the favourite BFT spawning areas, as indicated by their preference for waters with a salinity range of 36.9 to 37.7 (Alemany et al. 2010). It is unclear whether BFT adults detect salinity gradients or other processes

associated with the front, but the preference for specific salinity values does not seem to be related to specific physiological constrain but for areas were different water masses generates retention eddies and filaments (Alvarez-Berastegui et al. 2016).

### Chlorophyll-a

It has been consistently observed that BFT spawning grounds are often located in areas with low chlorophyll-a (CHLa) concentrations (Muhling et al. 2011, Koched et al. 2013, Muhling et al. 2013, Llopiz and Hobday 2015). Additionally, in the Balearic Islands spawning grounds, chlorophyll-a can be used as indicator of the recent Atlantic waters, which are characterised by lower chlorophyll-a but also lower salinity values than the resident Mediterranean waters. This has allowed the use of chlorophyll-a as a substitute variable in spatial distribution models of spawning aggregations when salinity values are absent or to complement the information related to the spatial distribution of water masses (Alvarez-Berastegui et al., 2016).

### Sea surface temperature

The timing and spatial distribution of BFT spawning is directly influenced by sea surface temperature (SST) either in the Mediterranean Sea (García et al. 2005, Alemany et al. 2010) or in other spawning areas (Koched et al. 2013, Reglero et al. 2014). BFT typically starts spawning once SST reaches 19-20°C, with larvae displaying a preference for waters between 23 and 28°C (Muhling et al. 2013). Furthermore, SST may also initiate gonadal development in adult BFT, as suggested by an increase in the gonadosomatic index of mature individuals upon reaching the Balearic Sea (Medina et al. 2002). High mean sea surface temperature may be also a requisite for embryo and larvae growth and survival, as they grow faster with increasing temperature, decreasing their developmental time (Gordoa and Carreras 2014). These two processes, gonadosomatic development and larval survival, may have driven evolutionary constraints for the location of spawning areas (Ciannelli et al. 2015).

### Sea surface temperature: temporal gradient

The spawning activity of yellowfin tuna usually begins after a steady increase in SST (Margulies et al., 2007). This could also be the case for BFT (Heinisch et al. 2008). In fact, the temporal gradient of SST over the previous fifteen days was useful for predicting BFT spawning ground around the Balearic Islands. As water temperature increases quite fast during early summer, 15 days and 7 days temporal gradients were considered (Álvarez-Berastegui et al., 2016).

### Fronts and mesoscale activity

The spawning ecology of BFT is also linked to mesoscale activity associated to the main fronts during the spawning season (Alemany et al., 2010). The highest abundances of young larvae, indicating proximity to spawning aggregations, are associated with low–medium kinetic energy values near the front, (Alvarez-Berastegui et al. 2016). The spatial gradients of salinity have been shown to be a good predictor of spawning grounds, especially when included in spatial models as interaction effect with salinity (Alvarez-Berastegui et al. 2014). In the western Mediterranean, the main frontal area is formed when recent Atlantic waters enter in the Mediterranean through the strait of Gibraltar and reach the Balearic Sea, mixing with saltier

resident waters (Balbín et al. 2014). Hence, the spatial gradients of SST (gradsst) and SSS (gradsal) were considered, as proxies of fronts between water masses. However, it has to be borne in mind that in the Mediterranean SST fronts generally weaken or disappear during the summer period. Furthermore, mixed layer depth (mld) and finite size Lyapunov exponents (fsle) are used as proxies of mesoscale mixing activity. Finite size Lyapunov exponents are a proxy of mixing activity that have been shown to be related with BFT larval abundance in the Balearic Islands (Díaz-Barroso et al. 2022). They are based on the separation rate of a pair of particle trajectories and measure the effects of transport and mixing mechanisms of water masses. The largest Lyapunov values occur along characteristic lines, called Lagrangian coherent structures (LCSs; Hernández-Carrasco et al. 2011, Bettencourt et al. 2012), which act as transport barriers, allowing a proper identification of fronts, eddies and filaments. Since LCSs cannot be crossed by particle trajectories, these structures strongly constrain and determine fluid motion, helping to analyse from a quantitative perspective how ocean transport is organized. Finite size Lyapunov exponents expressing sea surface mixing activity related with BFT larval presences in the western Mediterranean (Díaz-Barroso et al., 2022), being a more powerful explanatory variable than salinity.

## Methods

### Overview

Abundance was estimated in two ways, design-based estimation based on the survey coverage and estimated detection probability and secondly by using the estimated detection probabilities to produce estimated detection adjusted abundance and biomass estimates. In addition, a spatial model was fitted to detection adjusted estimated biomasses to detect potential drivers of biomass.

### Aerial surveys data (2017-2022)

Table 1 summarises the timing, number of transect, company and airplane type used for BFT surveys in 2017-2022. Not all blocks were surveyed each year.

Table 1. Summary of blocks and survey design.

Year	Block	Dates	Number of transects	Company	Airplane
<b>2017</b>	A	30 May – 26 June	29	Airmed	Partenavia-p68
	C	30 May – 14 June	25	Unimar	Partenavia-p68
	E	30 May – 01 July	30	Airmed	Partenavia-p68
<b>2018</b>	A	31 May – 28 June	36	Airmed	Partenavia-p68
	C	28 May – 16 June	25	Unimar	Partenavia-p68
	E	31 May – 21 June	40	Unimar	Partenavia-p68
<b>2019</b>	A	28 May – 28 June	30	Airmed	Partenavia-p68
	C	03 June – 16 June	23	Unimar	Partenavia-p68
	E	01 June – 04 July	40	ActionAir	Cessna 337

<b>2021</b>	A	04 June – 05 July	28	ActionAir	Cessna 337
<b>2022</b>	A	07 June – 27 June	30	Air Perigord	Cessna 337
	C	12 June – 20 June	25	Aerial Banners	Partenavia-p68
	E	22 June – 04 July	30	Aerial Banners	Partenavia-p68

## Extracting environmental covariates

### Data sources

Data sources included satellite and hydrodynamic models obtained from Copernicus, NOAA and IMEDEA-UIB-CSIC (Table 2). Derived variables were computed from the primary variables (Table 2). Repositories for the areas and time frames of interest were created locally through ftp connection. The software for the data extraction and calculation of the derived variables was developed in R language (R Core Team 2022), using the following packages: ncdf4 (Pierce and Pierce 2019), raster (Hijmans 2018), sf (Pebesma and Bivand 2023) and mgcv (Wood 2006).

### Spatial resolutions

The spatial resolutions for environmental variables were selected to accommodate spatial scales that have previously been found to have some influence on BFT larval habitat (Álvarez et al., 2016; Díaz-Barroso et al., 2022). Hence, selected spatial resolution for data extraction were i.e., 1.0°, 0.625°, 0.4° and 0.25° decimal degrees. Environmental data were retrieved for each sample observation, centred at segment locations (MidLon, MidLat), and calculated as the mean of the variable in a delta radius ( $\frac{1}{2}$  the spatial resolution). The acronym of each variable (see Table 1) has an added subscript referring to this delta radius i.e., \_d05, \_d03125, \_d02 and \_d0125 respectively (e.g., CHL\_d05, sst\_d05). Bathymetry (bathy) was retrieved using the library marmap (Pante and Simon-Bouhet 2013) at the lowest spatial resolution available i.e., 0.0042°x0.0042° decimal degrees, and at 0.07°x0.07° (Figure 2).

Table 2. Retrieved and derived variables for the segments collected from aerial surveys. Spatial resolution (Spat. Res.) is in decimal degrees.

Variable	Acronym	Units	Temporal scale	Spat. Res. product	Type of data	Institution	Spat. Res. source	Source
<b>Bathymetry</b>	bathy	m	-	0.017°	Model	NOAA 2022	0.0042°	1
<b>Slope</b>	slope	degrees	-	0.07°	Derived	CN-IEO-CSIC	0.0042°	-
<b>Slope categorical variable</b>	catslope	-	-	0.6°	Derived	CN-IEO-CSIC	0.0042°	-
<b>Day of the year</b>	doy	-	daily	-	Derived	CN-IEO-CSIC	-	-
<b>Presence absence of BFT</b>	pa	-	-	At segment locations	Derived	CN-IEO-CSIC	-	-
<b>Sea surface chlorophyll</b>	CHL	mg m <sup>-3</sup>	daily	1.0°, 0.625°	Satellite	Copernicus	0.042°	2

				0.4° & 0.25°				
<b>Sea surface temperature</b>	sst	°C	daily	1.0°, 0.625°, 0.4° & 0.25°	Satellite	Copernicus	0.05°	3
<b>Sea surface temperature –temporal gradient</b>	sst7dgrad	Δ°C	7 day	1.0°, 0.625°, 0.4° & 0.25°	Derived	CN-IEO-CSIC	0.05°	-
	sst15dgrad	Δ°C	15 day	1.0°, 0.625°, 0.4° & 0.25°	Derived	CN-IEO-CSIC	0.05°	-
<b>Residual temperature</b>	restemp	Δ°C	daily	1.0°, 0.625°, 0.4° & 0.25°	Derived - Model	CN-IEO-CSIC	0.05°	-
<b>Mixed layer depth</b>	mld	m	daily	1.0°, 0.625°, 0.4° & 0.25°	Model	Copernicus	0.042°	4, 5
<b>Sea surface salinity</b>	sal	-	daily	1.0°, 0.625°, 0.4° & 0.25°	Model	Copernicus	0.042°	4,5
<b>Finite size Lyapunov exponents</b>	fsle	day <sup>-1</sup>	daily	1.0°, 0.625°, 0.4° & 0.25°	Derived-Model	IMEDEA-UIB-CSIC	0.016°	6
<b>Sea surface temperature – spatial gradient</b>	gradsst	Δ°C	daily	0.5°	Derived	CN-IEO-CSIC	0.05°	-
<b>Sea surface temperature – spatial gradient</b>	gradsal	Δ	daily	0.5°	Derived	CN-IEO-CSIC	0.042°	-

1: <https://www.ncei.noaa.gov/products/etopo-global-relief-model>

2: [https://doi.org/10.25423/cmcc/medsea\\_multiyear\\_bgc\\_006\\_008\\_medbfm3](https://doi.org/10.25423/cmcc/medsea_multiyear_bgc_006_008_medbfm3)

3: <https://doi.org/10.48670/moi-00173>

4: For dates until 30/06/2021: [https://doi.org/10.25423/CMCC/MEDSEA\\_MULTIYEAR\\_PHY\\_006\\_004\\_E3R1](https://doi.org/10.25423/CMCC/MEDSEA_MULTIYEAR_PHY_006_004_E3R1)

5: For dates later than 30/06/2021:

[https://doi.org/10.25423/CMCC/MEDSEA\\_ANALYSISFORECAST\\_PHY\\_006\\_013\\_EAS7](https://doi.org/10.25423/CMCC/MEDSEA_ANALYSISFORECAST_PHY_006_013_EAS7)

6: Hernández-Carrasco *et al.*, 2011

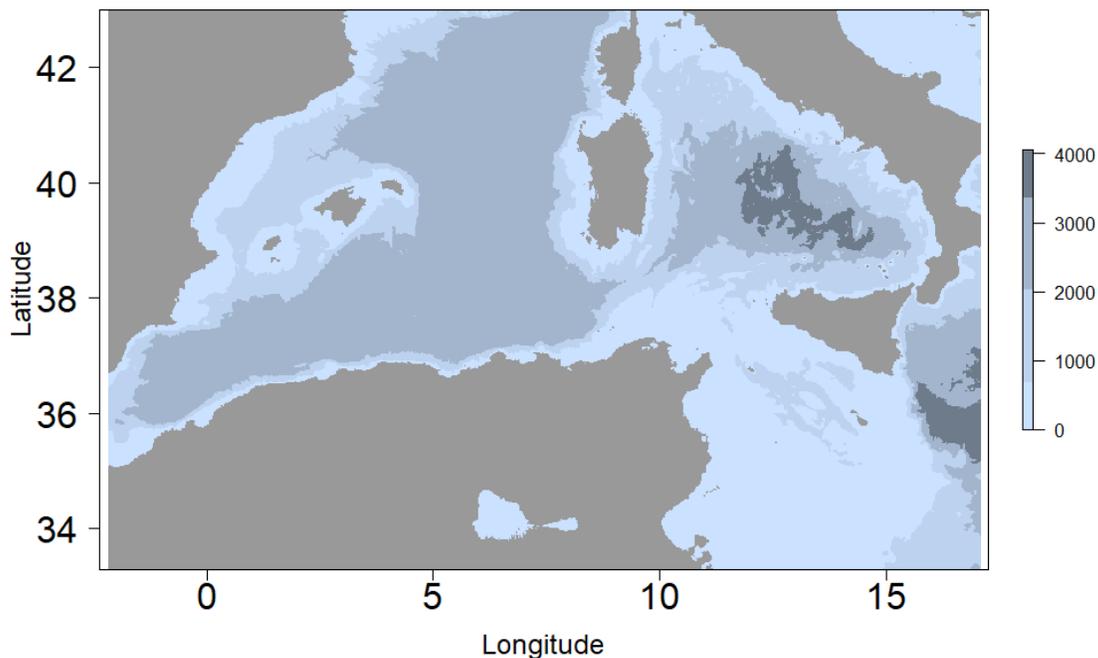


Figure 2. Map showing the extracted bathymetry (in meters) for the western Mediterranean.

### Primary and derived environmental variables

Bathymetry (bathy), sea surface chlorophyll (CHL), sea surface temperature (sst), mixed layer depth (mld), sea surface salinity (sal) and finite size Lyapunov exponents (flse) were retrieved for each segment of the aerial surveys from different data sources, i.e. satellite, oceanographic models or dispersion models (Table 2).

### Data processing and derived variables

Derived environmental variables were computed from the extracted primary data. The presence-absence of BFT (pa) was created was computed from the number of individuals (NumIndividuals) by providing a 0 to any observation with no detections and 1 to observations with BFT detections. The day of the year (doy) was computed from the variable Date using the library lubridate (Grolemund and Wickham 2011). Slope was calculated from extracted bathymetry at  $0.07^\circ \times 0.07^\circ$  spatial resolution and using the function terrain from the raster package (Hijmans 2022) (Figure 3). Additionally, a binary variable for slope (catslope) was created to identify observations that are on the shelf-break (1: slope  $\geq 2$ ) and areas that are outside the shelf-break (0: slope  $< 2$ ), after upscaling the variable slope to spatial resolution of  $0.6^\circ \times 0.6^\circ$  performed using a bilinear interpolation by a factor of nine (Figure 4).

Temporal gradients for sea surface temperature were computed as the mean difference of temperature at each cell of the grid at 7 and 15-day intervals following Álvarez-Berastegui et al., 2016. Residual temperature (restemp) was estimated at each spawning area (A, C, E), as the

residuals of a generalised additive model (fitted by restricted maximum likelihood) relating SST at each spatial resolution with the Julian day, to remove the effect of temperature increase along the season.

Spatial gradients for sea surface temperature and for salinity were computed as the maximum absolute difference between the mean hydrographic variable at the centre polygon and each of the eight surrounding polygons standardized to distance (Worm et al. 2005).

Finite size Lyapunov exponents were post-processed to avoid NAs corresponding to particles that ended up at the coast or that remained in the same place for more than 180 days ( $FSLE < 0.006$ , 1/days). This NA values were set to 0.006, which is the minimum detection limit.

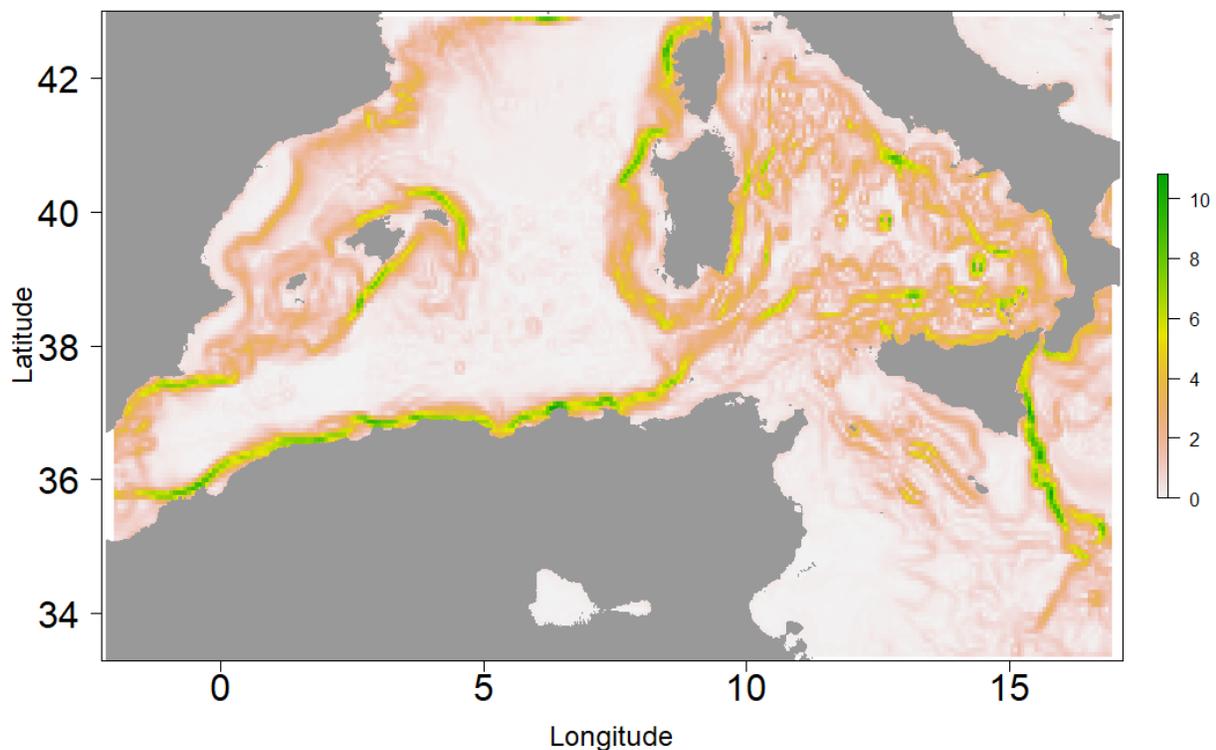


Figure 3. Map showing the slope variable (in radians) generated from the extracted bathymetry for the western Mediterranean.

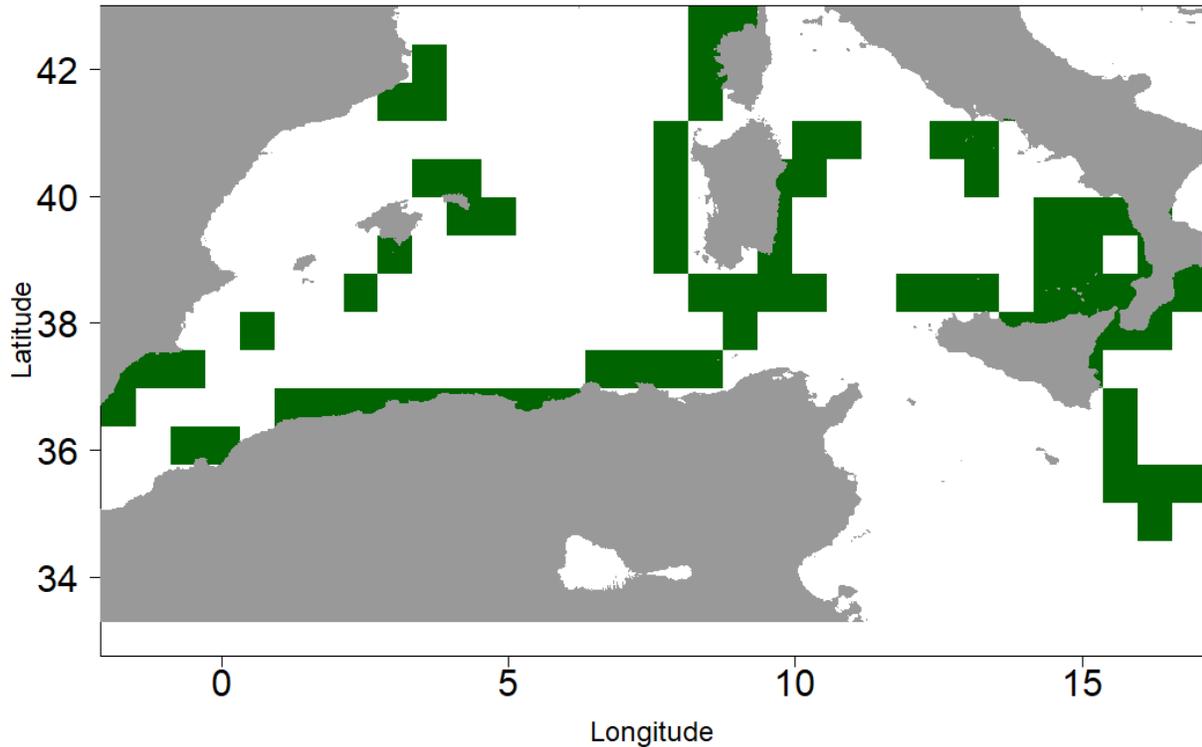


Figure 4. Categorical value of slope (catslope) for the western Mediterranean. In green (category 1) for shelf-break and in white (category 2) for non-shelf-break areas.

## Statistical methods - Distance sampling (DS)

### Survey design

We added 31 sightings from 2022 to the previously analyzed sightings data (Chudzinska et al. 2021, Chudzinska et al. 2022) (Table 3). Two sightings were excluded due to missing school sizes. Cluster size was obtained from TotalNumberPS regardless of composition in terms of size. However, no sightings with 100% small were provided with the data set.

Table 3. Number of sightings (within 1500m) per year and block. '-' indicate that no survey took place in a given block and year.

	2017	2018	2019	2021	2022
A	22 (18)	29 (24)	20 (19)	8 (7)	14 (8)
C	11 (7)	8 (8)	4 (4)	-	14 (11)
E	4 (4)	9 (7)	11 (6)	-	3 (3)
<b>Total</b>	<b>39 (29)</b>	<b>46 (39)</b>	<b>35 (29)</b>	<b>8 (7)</b>	<b>31 (22)</b>

As for previous surveys, we obtained perpendicular distances to the line for the 2022 sightings using the trigonometric relationship:

$$y_i = \tan((90 - \theta_i) \times 2 \times p_i/360) \times h_i$$

where  $y_i$  is the perpendicular distance between the line and the  $i$ th school,  $\theta_i$  is the declination angle and  $h_i$  the height of the airplane (Figure 5).

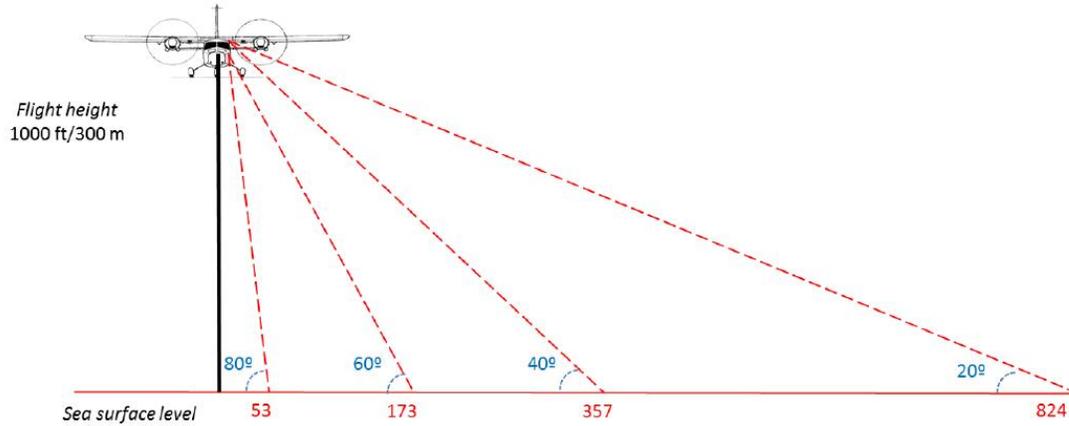


Figure 5. Correlation between the key declination angles and perpendicular distances at an altitude of  $h = 1000 \text{ ft} = 300 \text{ m}$  (Figure 5 from ICCAT survey protocol. Source: <https://www.iccat.int/>.)

A preliminary inspection of the distribution of perpendicular distances did not reveal any issues (Figure 6).

## 2017 - 2022

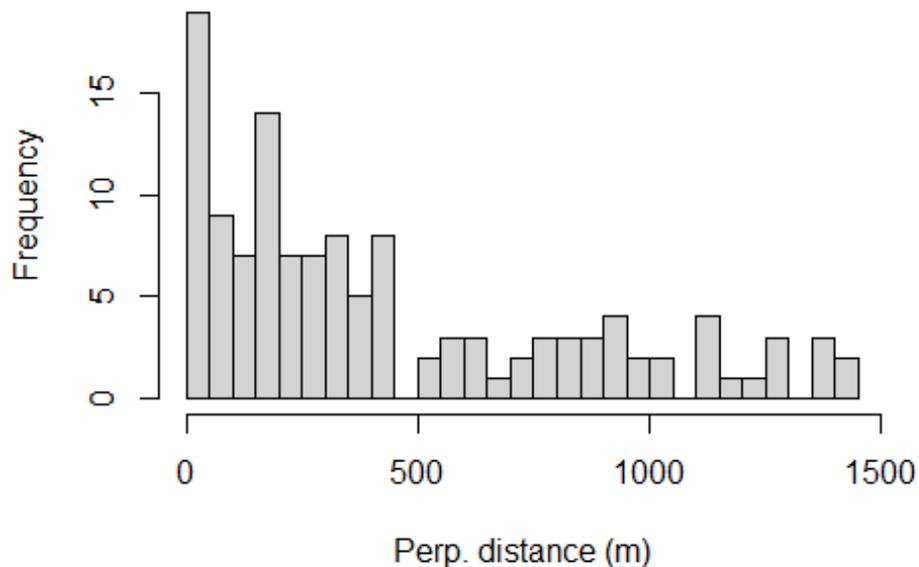


Figure 6. Histogram of detections within 1500m from the line.

### Fitting the detection function

Two critical assumptions of DS methods are that all schools on the transect centre line (i.e., at zero perpendicular distance) are detected with certainty and that distance measurements are exact. Given these assumptions, the distribution of perpendicular distances is used to model how the probability of detection decreases with increasing distance from the trackline.

Perpendicular distances were right truncated, where required, to avoid a long tail in the detection function, as well as left truncated, where required, to account for lower detection on the transect centre line. Left truncation is a common practice for aerial surveys, due to difficulties in searching directly underneath the plane, especially when the plane does not have a bubble window, which was not the case in the studied survey years. Perpendicular distances were truncated at 1,500 m. The choice of this truncation distance was based on visual inspection of fitted detection function, comparison with truncation distance used for previous years (2017-2021 models).

We used the packages *Distance* (Miller 2022) and *mrds* for analyzing the detection data in R version 4.2.0 (R Core Team 2022). We tested the fit of two key functions, the half normal (hn) and the hazard-rate (hr), and investigated whether adding covariates to the model in a multiple-covariate distance sampling (MCDS; e.g. Marques et al. 2007) approach would improve model fit. Here, six covariates which may affect the observers' ability to detect school were considered: the log of school size (*log-size*), the company conducting the survey (*company*),

*airplane*, *year*, *seastate* and a combination of survey block (A, C or E) with survey year (*block*) (Table 4). Model selection was based on minimum AIC values (Akaike 1987).

Table 4. Covariates considered for multiple-covariate distance sampling analyses.

Covariate	Description
log-size	Log of school size
company	Factor with five levels (Airmed, Unimar, ActionAir, Air Perigord, Aerial Banners)
airplane	Factor with two levels (Partenavia-P68, Cessna-337)
year	Factor with five levels (2017, 2018, 2019, 2021, 2022)
seastate	Continuous variable
block	Block-year combination

We conducted forwards model selection, starting with a model without covariates. In the next round we fitted single covariate models to either key function using the six available covariates. If any of the covariates improved the model, we investigated if adding a second covariate improved the fit in the next round. This process was repeated whereby one covariate was added to the best model from the previous round at a time until no improvement was achieved.

### Goodness of fit

Various options can be explored with function `Distance: gof_ds` which creates a quantile-quantile (Q-Q) plot and conducts various tests, including  $\chi^2$ , Kolmogorov-Smirnov and Cramer-von Mises tests, whereby large p-values indicate a good fit to the observed data (Miller et al. 2019, Laake et al. 2022).

The Q-Q plot shows how well the observed (or empirical) cumulative distribution function (ECDF) matches the fitted cumulative distribution function (FCDF) based on the best detection function model. Kolmogorov-Smirnov and Cramer-von Mises tests are based on the Q-Q plot and assess the deviations of ECDF  $\sim$  FCDF from a best fitting line (the  $y=x$  line, where the ECDF equals the FCDF for each data point in the sighting data). Hereby, the Kolmogorov-Smirnov test uses the largest vertical distance between a point and the  $y=x$  line as a statistic to test the null hypothesis that the samples (ECDF and FCDF) are from the same distribution and hence our model fits well. If the deviation between the  $y=xy=x$  line and the points is too large we reject the null hypothesis and say the model doesn't have a good fit.

By comparison, the Cramer-von Mises test uses all the differences between line and points and, while the null hypothesis is the same as for the Kolmogorov-Smirnov test, its test statistic is the sum of the deviations from each of the points to the line.

For the  $\chi^2$ -test, binning of distances was required as the sightings data are continuous. We used the default number of 20 equal-width bins. The test compares the number of observations in a given bin to the number predicted under the fitted detection function.

## Estimating group abundance

Detections and search effort were pooled within each survey to obtain encounter rates, and hence obtain estimates of density and abundance, by year (for 2017-2021 combined models). Estimates averaged overall surveys (weighted by survey effort) were also obtained.

The lengths of track lines were calculated from the recorded positions (i.e. latitude and longitude), when observers were on search effort. Only groups sighted when observers were on search effort and within 1,500 m were included in the analysis.

## Statistical methods - model-based inference

The count method of Hedley and Buckland (2004) was implemented to model the trend in spatial distribution in BFT. A common approach is to model the number of individuals in a small section of effort as a function of location and environmental descriptors. However, due to the nature of these data, where the range of group size can be from few to thousands (1 – 3,000 individuals in the data used in this model), a multi-step process was implemented:

- a model was fitted to the number of BFT groups (group encounter rate model)
- a model was fitted to group size (group size model)
- predictions from these two models were multiplied to produce a surface of BFT abundance.

For the purpose of this report, the same sightings, as used for DS, were considered: only sightings of schools consisting of adult individuals spotted by professional observers (for 2021 data).

Biomass estimation proceeded similarly to abundance estimation, first with the fitting a detection function followed by the fitting of a spatial model for the result and modified.

## Probability of detection

The probability of detection was estimated from the detection function modelled as described above.

## Modelling the number of groups

For model-based distance sampling analyses, the transect lines surveyed on effort were first divided into small sections (segments). The target length of segments was 10 km, but segments varied from this because of breaks in search effort. Sightings were attributed to the segments from which they were detected. Using the best fitting detection model (see above), the probability of detection  $p_r$ , was estimated for the  $r$ th school using the observed covariate values (if applicable) (Buckland et al. 2015).

The response variable  $N_i$  was the estimated number of groups in segment  $i$  with length  $l_i$ .  $N_i$  was calculated using a Horvitz-Thompson-type estimator (Horvitz and Thompson 1952):

$$\hat{N}_i = \sum_{r=1}^{R_i} \frac{1}{\hat{p}_r}$$

where  $R_i$  is the number of detected groups in segment  $i$ .

### Model specification

The estimated numbers of BFT groups  $\hat{N}_i$  per segment were used to estimate group abundance in the blocks of interest. This approach models spatial and temporal trends in the density and allows it to vary throughout the block and period of interest.

Counts are often modelled using a Poisson distribution; however, these data were over dispersed (i.e., more variable than expected for Poisson distributed data), and, therefore, we assumed a quasipoisson distribution, which allows a bit more variability of the counts.

The mean ( $\mu_i$ ) was modelled with location, habitat and temporal variables as candidate explanatory variables represented as follows.

$$\mu_i = \exp(\log_e(a_i) + \beta_0 + \sum_{j=1}^J \beta_j F_{ij} + \sum_{k=1}^K s_k(D_{ik}))$$

Where:

- $\log_e(a_i)$  is an offset term (a term with known regression coefficient) that corresponds to the area of each segment ( $a_i = 2wl_i$  where  $w$  is the truncation distance and  $l_i$  is the length of the  $i$ th segment),
- $\beta_0$  is an intercept,
- $\beta_j F_{ij}$  represent factor terms (e.g. year) with  $\beta_j$  representing the regression coefficients for the  $j$ th factor variable,
- $s_k(D_{ik})$  represent one dimensional smooth terms (e.g. depth)

Interaction terms can be added as two-dimensional smooths between two continuous covariates or between a continuous and a factor covariate which allows the smooth to vary for each factor level.

- $s_l(X_i, Y_i)$  represents a two-dimensional smooth term (determined for the  $i$ th segment) where at both  $X_i$  and  $Y_i$ .

The models were fitted using generalised additive models in the R package mgcv (Wood and Wood 2015).

The spatial model for biomass was generated in a similar manner with the original biomass estimates inflated by being divided by the estimated detection probability. These modified biomass estimates were in turn modelled spatially. However, in this case the distributional properties of the estimated abundances were quite awkward. For this reason, instead of being modelled as counts, the biomass data were modelled as densities in a zero-inflated model assuming a Poisson distribution of residuals for the non-zero component with a binomial presence absence component. Smooths were considered as before.

### Potential issues

We checked for collinearity and concurvity using the vif function from the car library which calculates variance-inflation and generalized variance-inflation factors (VIFs and GVIFs) (Fox and Weisberg 2018). VIFs can be used for comparing linear terms, GVIFs for factor or smooth terms (where the nonlinear equivalent to collinearity is called concurvity).

### Model selection

After eliminating collinear and concurvity covariates we fitted a full model with the remaining terms. Here, covariates were tested as smooth functions or, in specific cases, as interaction terms. The selection of interaction terms to be tested was based on discussion among co-authors. From this full model, we applied backwards model selection using p-values by checking if all model terms were significant. If any term had an associated p-value of greater than 0.05, we eliminated the term with the highest p-value and refitted the model. This process was continued until all model terms were significant. If an interaction term was non-significant, the main effects were tested and retained if significant.

### Group size modelling

Group (shoal) size could, like density of group numbers could vary over the range of the surveys therefore shoal size was modelled over the space and time of the surveys using the same methods as in the case of group number with the same consideration of variables except with the additional consideration of estimated distance to account for a potential bias that small size shoals would not be detected at greater distances. However, as the sample size was extremely small ( $n = 126$ ), interactions were not considered, only main effects. Again, model selection was backwards with a  $P < 0.05$  inclusion criterion.

Group sizes were modelled assuming a Gamma distribution of the residuals with an inverse gamma link function (as group size could not be zero).

$$\mu_i = 1 / (\beta_0 + \sum_{j=1}^J \beta_j F_{ij} + \sum_{k=1}^K s_k (D_{ik}))$$

where

- $\beta_0$  is an intercept,

- $\beta_j F_{ij}$  represent factor terms (e.g. year) with  $\beta_j$  representing the regression coefficients for the  $j$ th factor variable,
- $s_k(D_{ik})$  represent one dimensional smooth terms (e.g., depth)

The models were similarly fitted using generalised additive models in the R package mgcv (Wood and Wood 2015).

### Estimation of the total number of individuals

With spatial models for both group number and group size known then estimation of overall numbers and density per survey area could be undertaken by simply predicting the numbers for each model for each year for a single approximate optimum date from inspection of the data. However, this date could well vary between years and blocks. The prediction grid consisted of cells at 0.25° longitude and latitude intervals. Primary and derived environmental variables were retrieved for the variables selected in the final models and for the day with maximum mean predicted Bluefin tuna i.e., 31<sup>st</sup> of May of each year, and for the 12<sup>th</sup> of June, which is a secondary maximum of mean predicted Bluefin tuna number of groups.

### Estimation of uncertainty

Uncertainty was estimated by means of a non-parametric bootstrap. In the case of the group number and group size models for each iteration, new model parameters were generated based on the parameter estimates from the initial models along with their associated variance/covariance matrix.

## Results

### Distance sampling (DS)

#### Best fitting detection functions

Models tested are shown in Table 5. The same detection functions were used for abundance of groups and biomass. Including *log-size* as a covariate showed the best improvement of model fit in single-covariate models. Covariate *company* further improved the model when using the half-normal key function. Hence, the best model was the half normal with *log-size* and *company* as covariates (mod1722\_hn.size.company) and was used in the following to produce summaries, tables and plots.

Table 5. Models tested in forwards selection for group size where horizontal lines separate the different rounds of model fitting.  $\Delta$  AIC refers to the difference between a given model and best (underlined).

Models	key	Covariates	AIC	$\Delta$ AIC
mod1722_hn	hn	--	1774	45
mod1722_hr	hr	--	1773	44
mod1722_hn.company	hn	company	1790	61
mod1722_hn.plane	hn	plane	1790	61

<b>mod1722_hn.reg</b>	hn	block	1789	60
<b>mod1722_hn.seas</b>	hn	seastate	1790	61
<b>mod1722_hn.size</b>	hn	size	1732	3
<b>mod1722_hn.year</b>	hn	year	1797	68
<b>mod1722_hr.company</b>	hr	company	1762	33
<b>mod1722_hr.plane</b>	hr	plane	1772	43
<b>mod1722_hr.reg</b>	hr	block	1764	35
<b>mod1722_hr.seas</b>	hr	seastate	1774	45
<b>mod1722_hr.size</b>	hr	size	1743	14
<b>mod1722_hr.year</b>	hr	year	1774	45
<b><u>mod1722_hn.size.company</u></b>	<u>hn</u>	<u>size + company</u>	<u>1729</u>	<u>0</u>
<b>mod1722_hn.size.plane</b>	hn	size + plane	1734	5
<b>mod1722_hn.size.reg</b>	hn	size + reg	1735	6
<b>mod1722_hn.size.seas</b>	hn	size + seastate	1734	5
<b>mod1722_hn.size.year</b>	hn	size + year	1737	8
<b>mod1722_hr.size.company</b>	hr	size + company	1735	6
<b>mod1722_hr.size.plane</b>	hr	size + plane	1745	16
<b>mod1722_hr.size.reg</b>	hr	size + reg	1734	5
<b>mod1722_hr.size.seas</b>	hr	size + seastate	1745	15
<b>mod1722_hr.size.year</b>	hr	size + year	1746	17
<b>mod1722_hn.size.company.seas</b>	hn	size + company + seastate	1731	2
<b>mod1722_hr.size.company.seas</b>	hr	size + company + seastate	1736	7

### Goodness of fit

The fit of the best detection function model to the observed data was deemed adequate as judged by the Q-Q plot (Figure 7) and the three goodness of fit test statistics including the  $\chi^2$ -test (using 20 equally-spaced distance bins), the Kolmogorov-Smirnov test (with 100 bootstrap samples) and the Cramer-von Mises test (Table 6).

The histogram of detections (Figure 6) showed a relatively quick drop off in detection probabilities between 0 and 500 m (Figure 7). This was mostly driven by the detections with small school sizes, shown in the figure as those below the histogram line, i.e. lower than average detection probabilities. Most of these small school sizes were detected in Block A by company Airmed.

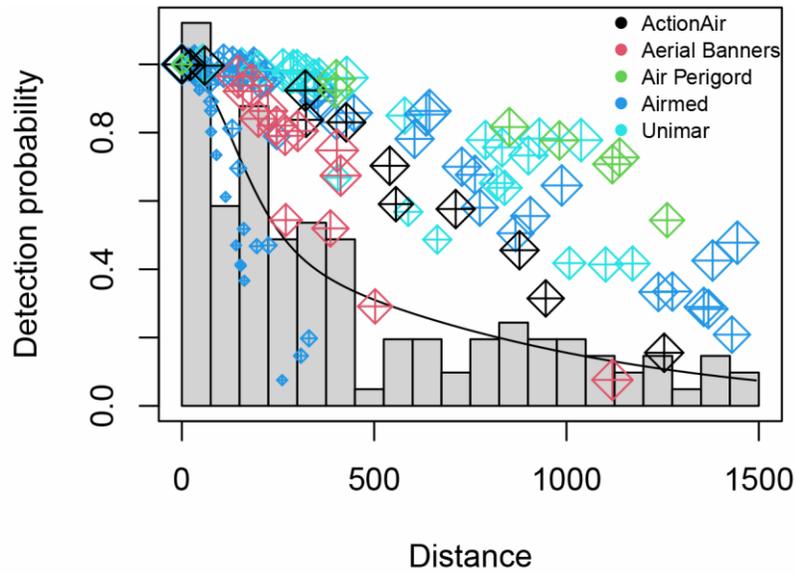


Figure 7. Histogram of observed distances, average detection function across all observations (histogram line) and detection probabilities of observed distances from best fitting model (dots) colour coded by company. Size of symbols were scaled to represent the natural log of school size.

Table 6. Goodness of fit tests and results conducted to assess fit of best detection function model.

Test	Test statistic	p-value
$\chi^2$ (13 degrees of freedom)	15.136	0.30
Kolmogorov-Smirnov	0.061	0.58
Cramer-von Mises	0.090	0.63

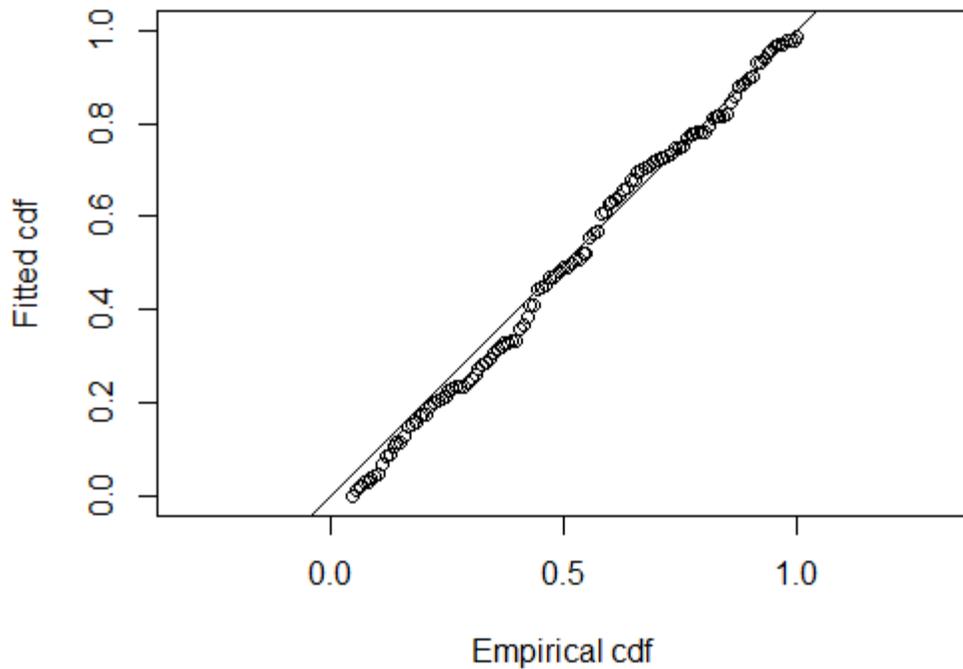


Figure 8. Q-Q plot for best fitting model showing the observed (empirical) cumulative distribution function (ECDF) against the fitted CDF (FCDF) along with a line of best fit ( $y=x$  line).

#### Summary of effort and estimates of encounter rate.

The largest effort per block was conducted in block E and lowest in block C (Table 7). However, most sightings were made in block A, resulting in the highest estimated encounter rates for block A with lowest CVs between 2017 and 2019. Encounter rate estimates were lowest in block E with higher than average CVs.

Table 7. Summary per block of area covered by the survey, effort conducted, number of groups encountered (n), number of transect lines (k), estimated encounter rate (ER) and its standard deviation (SE) and CV (CV).

Block	Area (km <sup>2</sup> )	Covered Area (km <sup>2</sup> )	Effort (km)	n	k	ER	SE	CV
A-2017	61837.1	14988.2	4996.1	18	52	0.003603	0.001048	0.29
A-2018	61837.1	18462.0	6154.0	24	62	0.003900	0.000858	0.22
A-2019	61837.1	16378.6	5459.5	19	50	0.003480	0.000774	0.22
A-2021	61837.1	18792.7	6264.2	7	46	0.001117	0.000591	0.53
A-2022	61837.1	16101.6	5367.2	8	38	0.001491	0.000554	0.37
C-2017	51821.2	14490.0	4830.0	7	25	0.001449	0.000469	0.32
C-2018	51821.2	14792.5	4930.8	8	25	0.001622	0.000646	0.40
C-2019	51821.2	14453.0	4817.7	4	23	0.000830	0.000470	0.57
C-2022	51821.2	14810.1	4936.7	11	25	0.002228	0.000775	0.35

E-2017	90101.5	19147.3	6382.4	4	41	0.000627	0.000302	0.48
E-2018	90101.5	26464.0	8821.3	7	47	0.000794	0.000395	0.50
E-2019	90101.5	24996.0	8332.0	6	48	0.000720	0.000322	0.45
E-2022	90101.5	19623.7	6541.2	3	32	0.000459	0.000331	0.72
Total	876876.3	233499.7	77833.2	126	514	0.001588	0.000159	0.10

Estimated numbers of groups  $N_G$  per block were highest again in block A between 2017 and 2019 which also had the lowest CVs. CVs for other blocks ranged between 39% and 73%. Estimated numbers of schools were lowest in block C in 2019 (Table 8).

Table 8. Number of groups encountered (n), estimated number of groups  $N_G$  per block along with its standard error (SE), coefficient of variation (CV), lower and upper 95% confidence limits (LCL and UCL).

Block	N	$N_G$	SE	CV	LCL	UCL
A-2017	18	430.6	159.1	0.37	210.9	879.4
A-2018	24	337.7	117.4	0.35	172.4	661.6
A-2019	19	307.8	86.0	0.28	178.2	531.6
A-2021	7	45.3	27.0	0.60	15.0	136.7
A-2022	8	45.0	24.7	0.55	16.3	124.6
C-2017	7	34.0	13.3	0.39	15.9	72.7
C-2018	8	47.9	23.4	0.49	19.0	120.7
C-2019	4	19.2	11.9	0.62	6.0	62.0
C-2022	11	105.4	42.6	0.40	48.1	230.9
E-2017	4	26.7	13.2	0.50	10.4	68.6
E-2018	7	31.6	17.1	0.54	11.5	86.9
E-2019	6	44.0	22.3	0.51	16.9	114.3
E-2022	3	55.9	40.7	0.73	14.9	209.8

The expected school sizes were lowest in block A, in particular during 2017 – 2019 with estimates of less than 250 individuals per school (Table 9). Blocks C and E generally had high school sizes. For block C, all but the 2018 estimate were larger than 1000 individuals. For block E, estimates declined from a high in 2017 to a low in 2022.

Table 9. Expected school size per block ( $E[size]$ ) and standard error (SE)

Block	$E[size]$	SE
A-2017	111.3	60.0
A-2018	231.5	100.5
A-2019	241.1	96.7
A-2021	979.8	190.2
A-2022	572.3	188.7
C-2017	1297.6	316.9
C-2018	766.9	345.5
C-2019	1326.3	406.1
C-2022	1592.6	176.8
E-2017	1581.1	324.1

E-2018	1241.0	357.7
E-2019	709.8	64.4
E-2022	408.7	120.4
Total	444.2	84.6

For block A, estimated number of individuals were highest in 2018 and declined steadily to 2022. For block C, the estimated number of individuals declined between 2017 and 2021, however, came to an overall high in 2022. For block E, the estimated number of individuals declined steadily between 2017 and 2022. Overall, estimated numbers of individuals  $N$  were highest in block C in 2022. Block C also showed largest increase in abundance in comparison to previous years. Abundance estimates in blocks A and E are comparable to the last surveys conducted in these blocks. CVs for the individual blocks were overall high – lowest for block A with 31% but generally over 40% otherwise and up to 83% in block E in 2022 (Table 10, Figure 9). Figure shows comparison between current estimates and estimates of the abundance from the previous reports. The current results are comparable with the previous estimates.

Table 10. Estimated number of individuals  $N$  (in thousands) per block (block and year) with standard errors and 95% CIs. The same columns with 'str' apply to estimates reported in [Chudzinska et al. 2021](#), and [Chudzinska et al. 2022](#). '-' indicates that estimates previous to this study were not provided.

Label	N	SE	CV	LCL	UCL	N_str	CV_str	LCL_str	UCL_str
A-2017	47.93	20.56	0.43	21.06	109.11	49.92	0.44	21.82	114.2
A-2018	78.18	23.94	0.31	43.10	141.81	81.6	0.31	45.28	147.1
A-2019	74.22	29.41	0.4	34.54	159.48	75.02	0.38	36.71	153.3
A-2021	44.33	24.86	0.56	15.58	126.16	26.11	0.54	9.59	71.13
A-2022	25.78	11.87	0.46	10.76	61.75	-	-	-	-
C-2017	44.11	18.00	0.41	19.87	97.90	44.89	0.44	19.52	103.2
C-2018	36.77	21.03	0.57	12.38	109.22	37.38	0.53	13.72	101.8
C-2019	25.53	14.77	0.58	8.46	77.05	25.98	0.61	8.43	80.06
C-2022	167.82	70.57	0.42	74.22	379.47	-	-	-	-
E-2017	42.16	21.87	0.52	15.76	112.78	44.1	0.54	16.38	118.7
E-2018	39.24	23.01	0.59	13.20	116.62	40.1	0.47	16.63	96.71
E-2019	31.21	15.65	0.5	12.15	80.18	17.93	0.51	6.96	45.68
E-2022	22.85	19.05	0.83	5.23	99.82	-	-	-	-

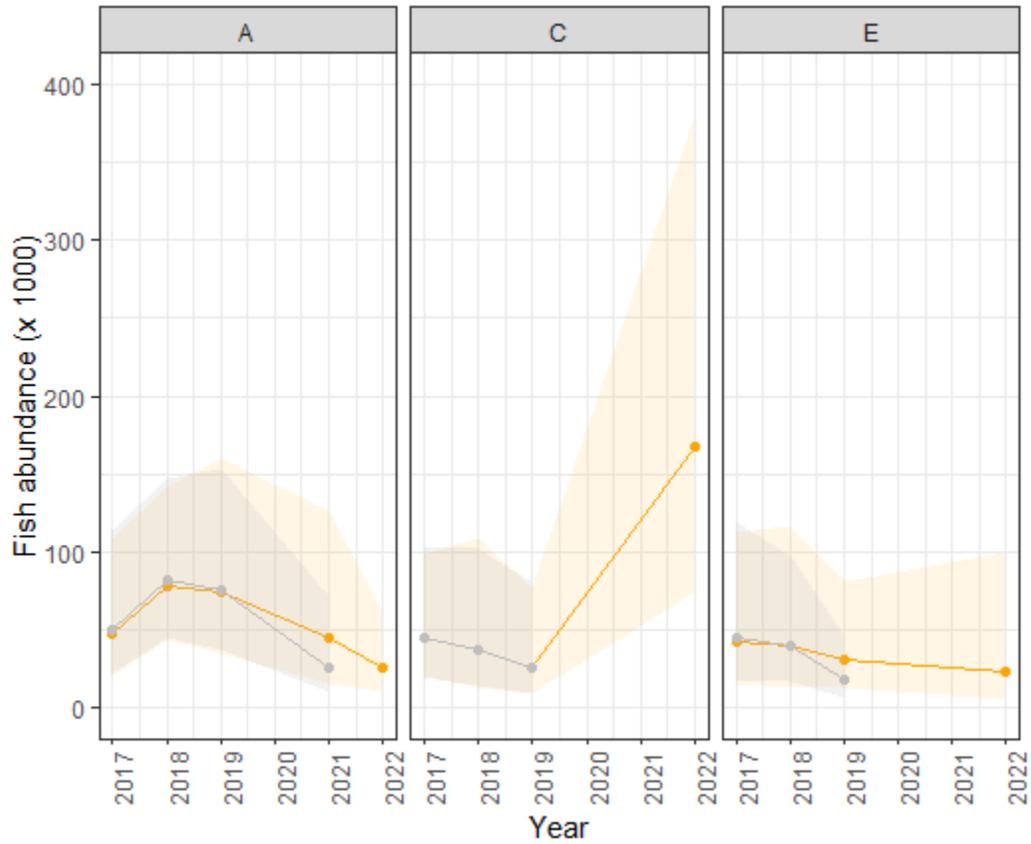


Figure 9. Estimated abundance of BFT for surveyed years and blocks. Orange colour show estimates from this study: dots show mean values and ribbon show upper and lower confidence limits of the 95% confidence interval. Grey colour shows estimates from the previous reports (Chudzinska et al. 2021, and Chudzinska et al. 2022) for comparison ('strict update').

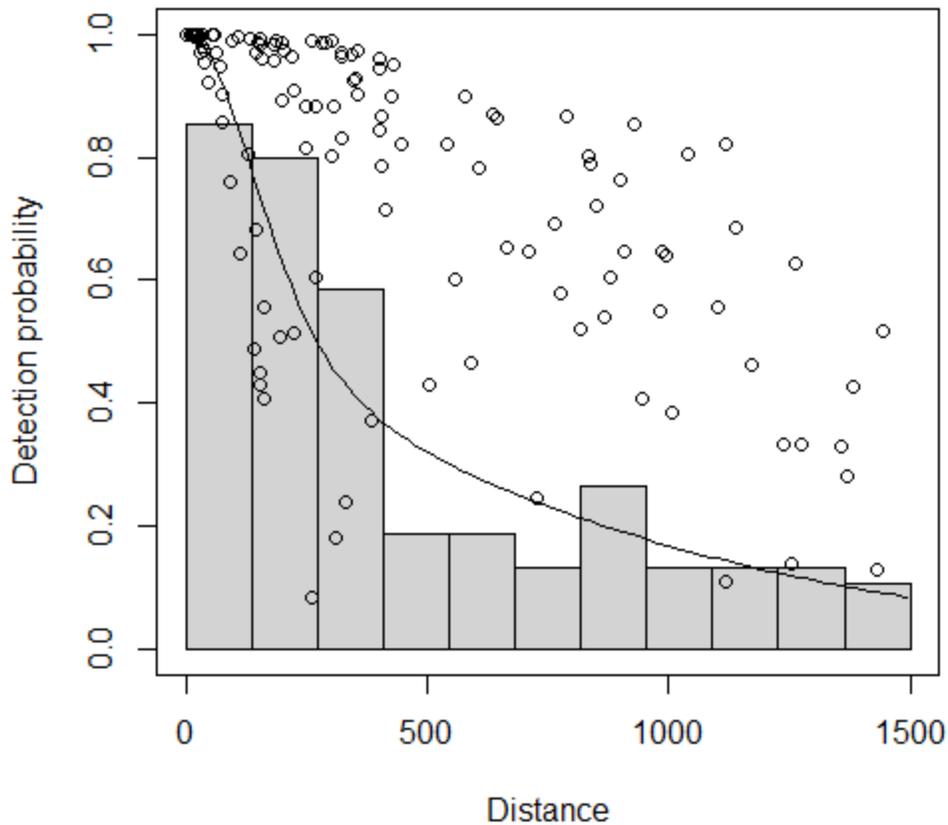


Figure 10. Histogram of observed distances, average detection function across all observations (histogram line) and detection probabilities of observed distances from best fitting model (dots) with log (Biomass) the sole covariate apart from distance.

The biomass estimates for the previous years were based on a detection function (see Figure 10) with distance and log (Biomass). The biomass estimates are presented in Table 11 and Figure 11. The results are comparable with the estimates from the previous reports but the estimate for biomass is higher for block A in 2021. This appears to be because G block observations were not used in constructing the detection function. For reasons of backwards compatibility with previous strict update results, results based on a detection function

including the Block G data were included in Appendix 2. The biomass in block E is comparable to the estimates from 2019 when the last survey took place in that block.

Table 11. Estimated biomass (B, in tons) per block along with its standard error (SE), coefficient of variation (CV), lower and upper 95% confidence limits (LCL and UCL). The same columns with 'str' apply to estimates reported in [Chudzinska et al. 2021](#), and [Chudzinska et al. 2022](#). '-' indicates that estimates previous to this study were not provided.

	B	CV	LCL	UCL	B_str	CV_str	LCL_str	UCL_str
A-2017	8927	0.43	3879	20546	8001	0.45	3436	18634
A-2018	14857	0.3	8227	26828	13345	0.31	7352	24222
A-2019	13268	0.39	6202	28387	11548	0.38	5619	23734
A-2021	7667	0.53	2832	20760	4714	0.53	1750	12696
A-2022	11903	0.43	5129	27627	-	-	-	-
C-2017	7691	0.38	3620	16338	6749	0.43	2981	15280
C-2018	5822	0.55	2022	16764	5069	0.54	1846	13920
C-2019	3549	0.56	1198	10513	3072	0.62	977	9652
C-2022	6878	0.41	3065	15436	-	-	-	-
E-2017	5639	0.58	1899	16750	5884	0.6	1981	17483
E-2018	4423	0.54	1601	12215	3735	0.47	1538	9067
E-2019	3063	0.46	1268	7396	2023	0.5	797	5188
E-2022	1346	0.74	352	5149	-	-	-	-

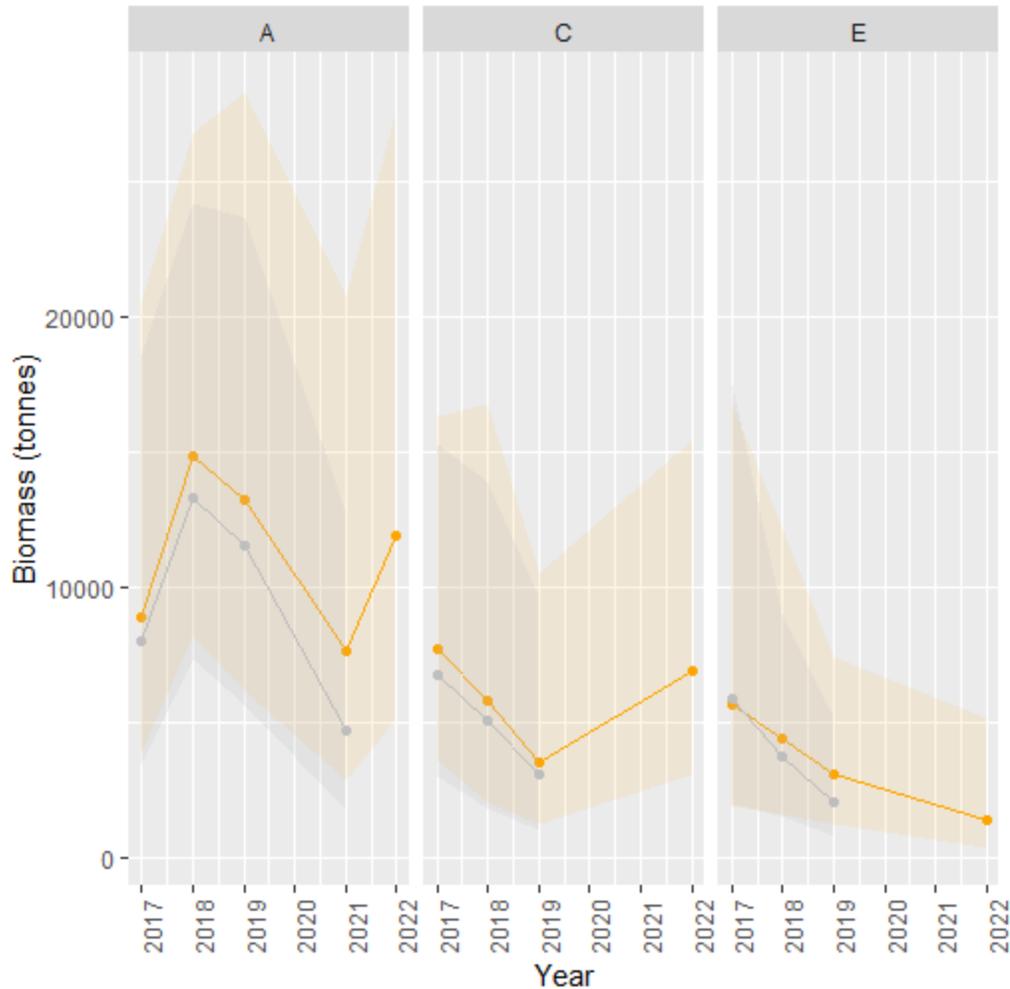


Figure 11. Estimated biomass (in tons) of BFT for surveyed years and blocks. Orange colour show estimates from this study: dots show mean values and ribbon show upper and lower confidence limits of the 95% confidence interval. Grey colour shows estimates from the previous reports (Chudzinska et al. 2021, and Chudzinska et al. 2022) for comparison ('strict update').

## Exploration of candidate explanatory variables

In this section we explored the relationship between the response variable for the model  $N_i$  (number of groups per segment adjusted for imperfect detection) and each of the covariates. The available candidate explanatory variables were divided into 11 groups.

### Group 1: bathy, catslope, Julian

Covariate *bathy* showed higher responses in areas where this covariate was less than 2000 m (Figure 12). Mean responses were increasing for *bathy* values up to 2000 m and declined thereafter. Responses varied similarly for both levels of *catslope*. However, mean response was

higher for *catslope* = 1. *Julian* had the highest responses and larger variability in the central part of observed values (i.e. in the middle of the survey period). However, highest mean response was observed on *Julian* = 151.

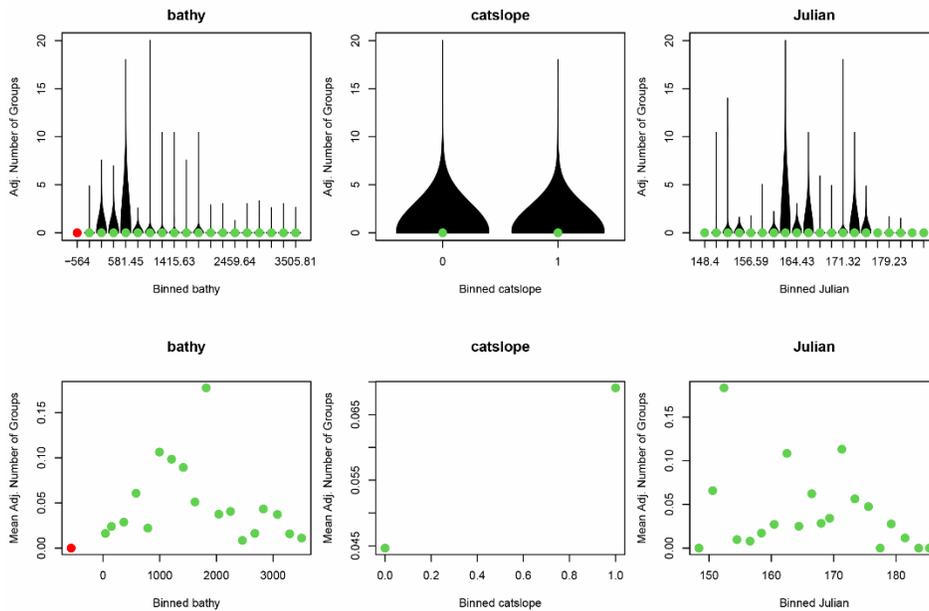


Figure 12. Relationship between the response and variables bathy, catslope and Julian and. Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

## Group 2: Slope with varying spatial resolution

All slope variables had very few observations in the higher ranges of these covariates (Figure 13). Highest variability in the responses was observed in the lower areas of these covariates. Another feature shared between these covariates was that mean responses increased through the lower half of observed values and generally decreased in the higher half. However, highest mean response for slope was observed in the second highest bin. This was due to one out of seven observations with a non-zero response in this bin.

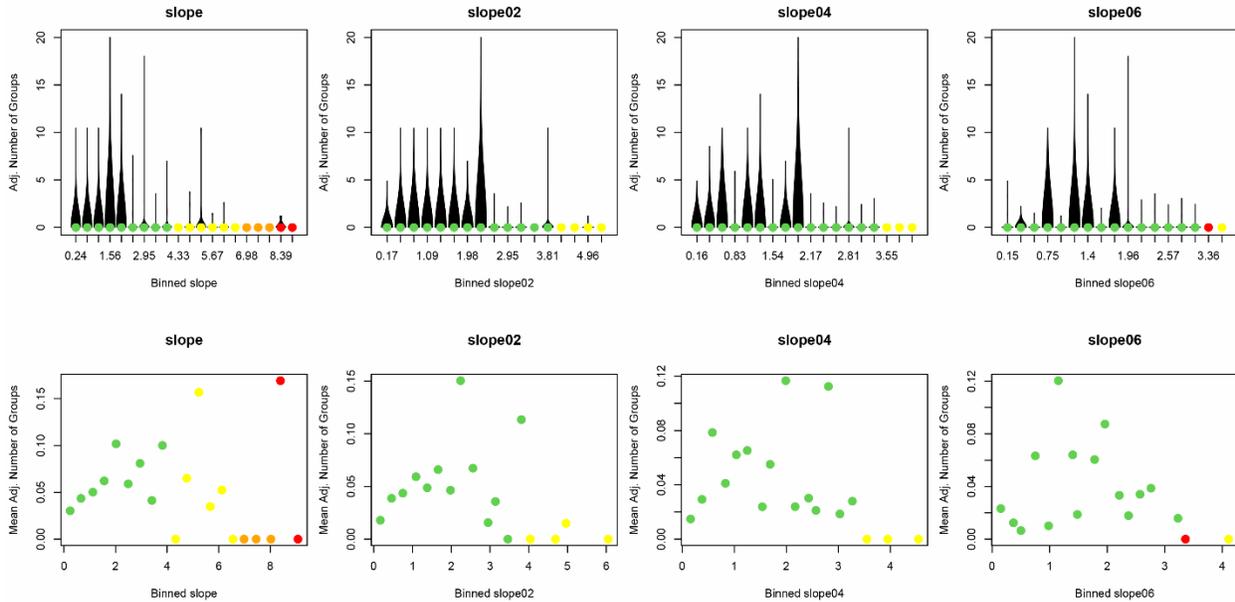


Figure 13. Relationship between variables related to slope and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group3: Chlorophyll – daily product

For all *chlorophyll* variables, the majority of observations were within the lowest two bin (Figure 14). Mean responses were generally highest in these two bins as well. However, for *CHL\_d05* showed a highest mean response in bin 3.

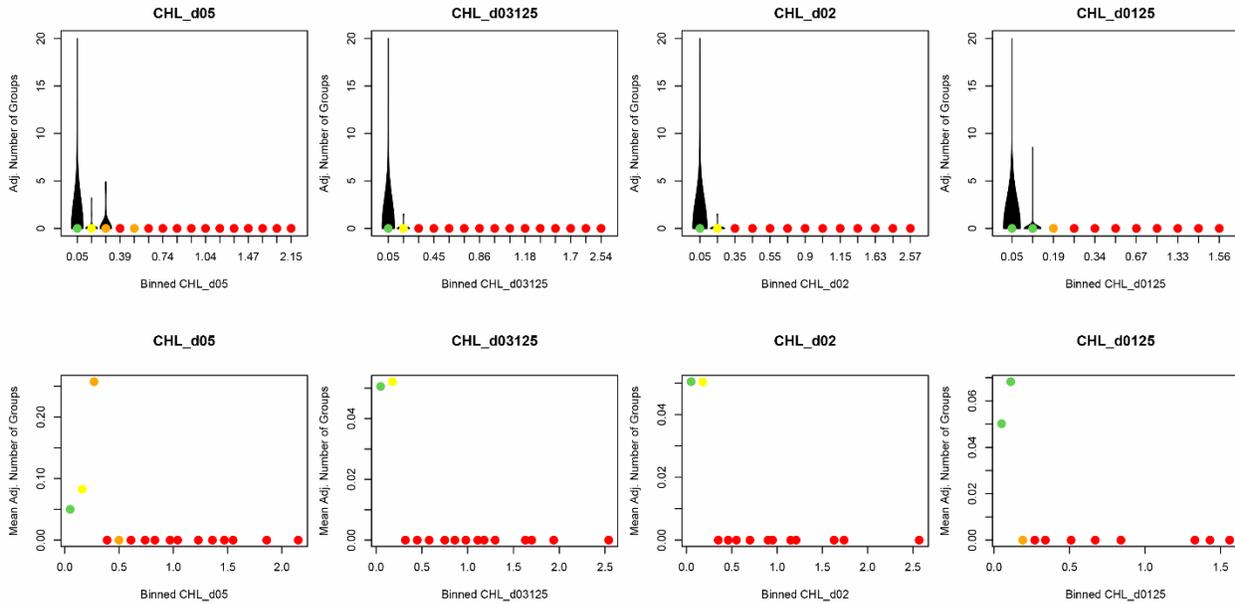


Figure 14. Relationship between variables related to chlorophyll and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

#### Group 4: Sea-surface temperature – daily product

Covariates related to sea-surface temperature showed highest variability in responses in the central areas of observed values, i.e. between approximately 21 and 24 °C (Figure 15). Outer blocks of these covariates had few observations. Highest responses were observed around 24 °C while highest mean responses were observed between 20 and 21 °C, depending on the covariate.

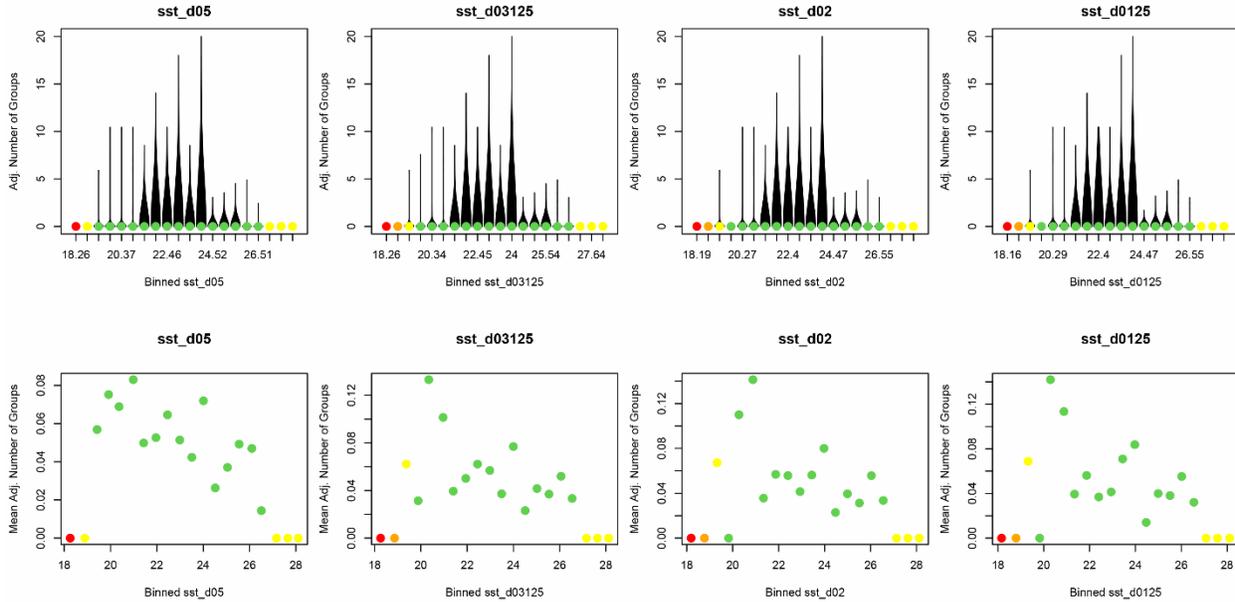


Figure 15. Relationship between variables related to daily average sea-surface temperature and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group 5: Sea-surface temperature – 7-day temporal gradient

Covariates related to a 7-day temporal gradient in sea-surface temperature also had the highest variability in the response in the central parts of the observed covariates and few observations in the lowest or highest bins (Figure 16). However, each covariate had the highest mean response in the lowest category due to two non-zero response segments in this bin.

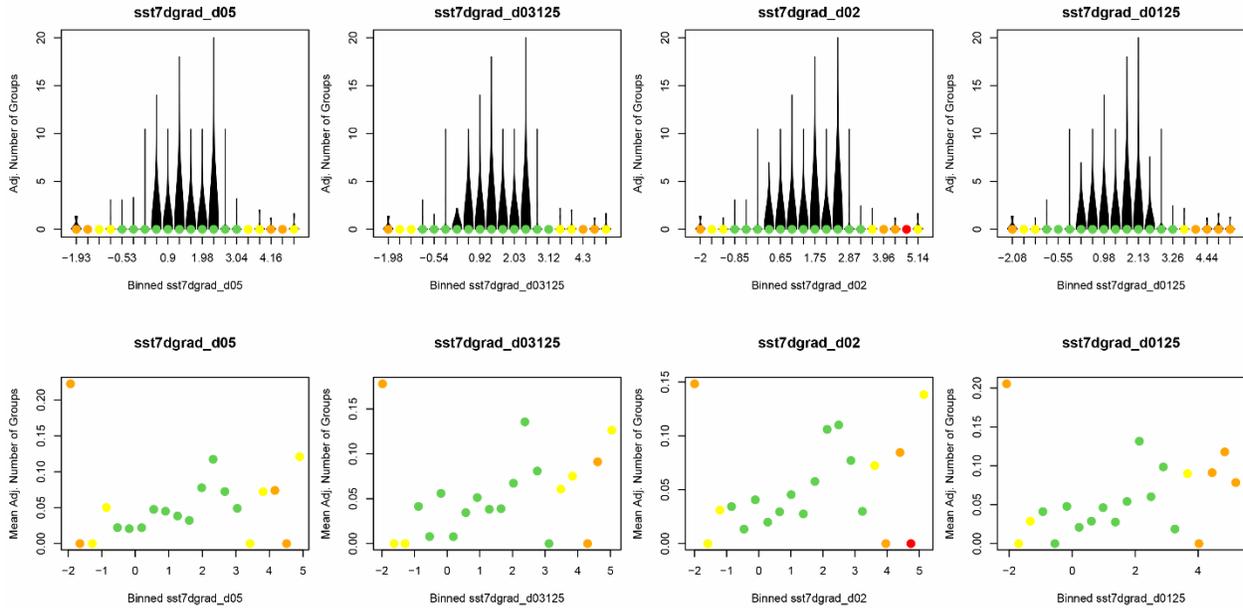


Figure 16. Relationship between variables related to 7-day gradient in sea-surface temperature and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group 6: Sea-surface temperature – 15-day temporal gradient

Covariates related to the 15-day temporal gradient showed strong variability in the lower half of the observed ranges in these covariates (Figure 17). The outer most bins, again, had few observations, in particular toward the upper end. This resulted in three of the four covariates in this group to have the highest mean response in one of the highest bins.

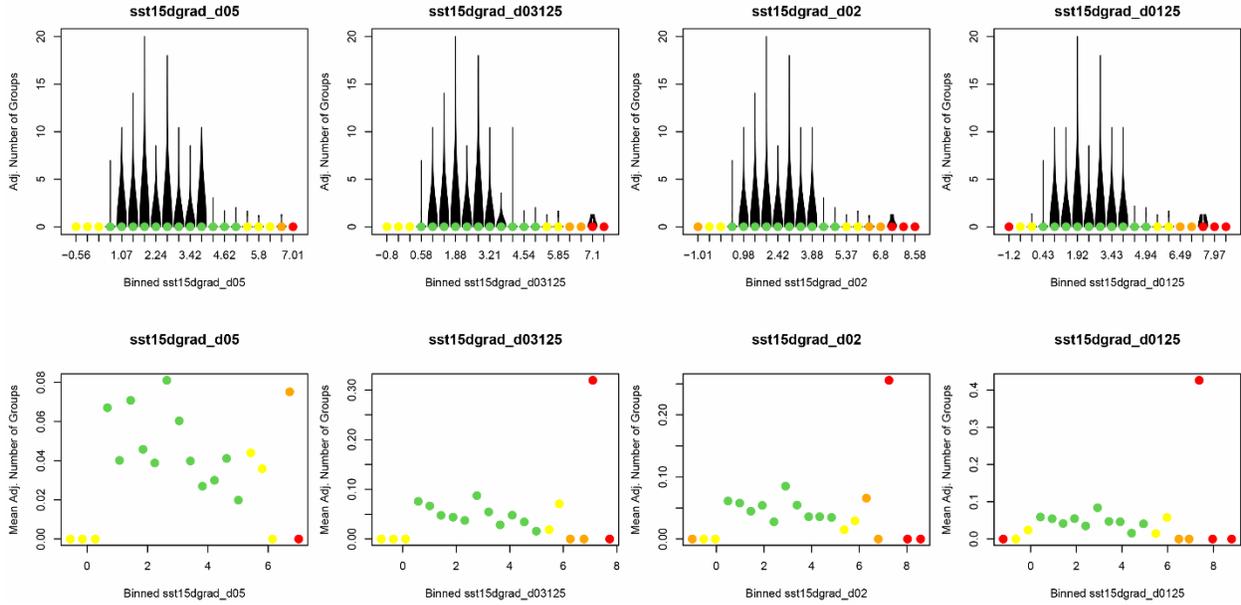


Figure 17. Relationship between variables related to 15-day gradient in sea-surface temperature and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group 7: Mixed layer depth – daily product

Most observations for covariates in this group fell within two, three or four bins which also contained the majority of non-zero responses (Figure 18). Non-zero means were also constrained to few categories with on distinct pattern.

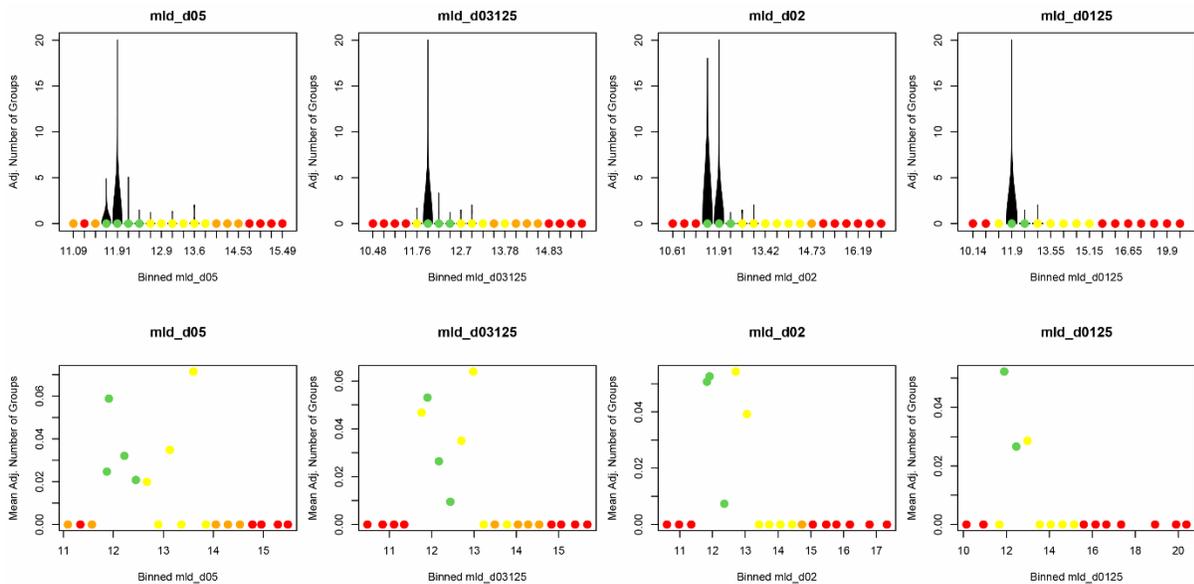


Figure 18. Relationship between variables related to mixed layer depth and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group 8: Salinity – daily product

Covariates in the salinity group showed most variability in the response in the central part of these covariates (Figure 19). However, mean responses varied with no distinct pattern.

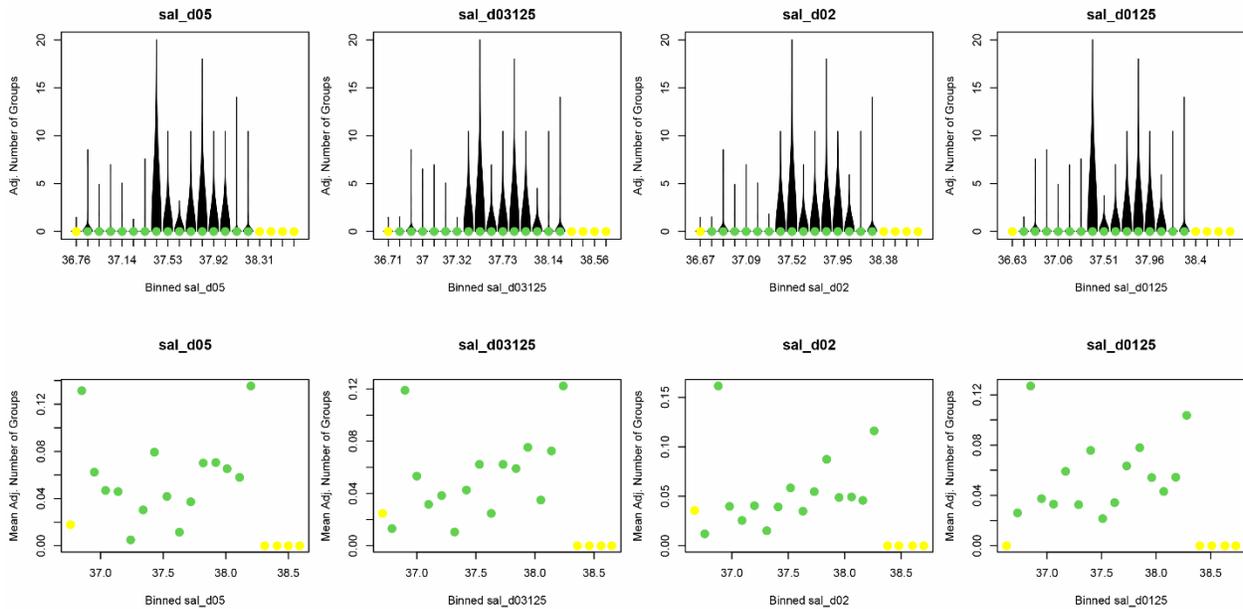


Figure 19. Relationship between variables related to salinity and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group 9: Finite size Lyapunov exponents – daily product

Depending on the spatial resolution, covariates in this group either showed the highest variability in responses in the central bins of the covariates (*fsle\_d05* with the highest spatial resolution) (Figure 20). As the spatial resolution decreased, the bins with the highest variability shifted toward the lower bins. A similar pattern was evident for the mean responses. For covariate *fsle\_d03125*, a single non-zero response segment out of five segments in the second highest bin produced the maximum mean response out of all bins. These covariates had 1728 NAs, hence, we excluded this group from model selection.

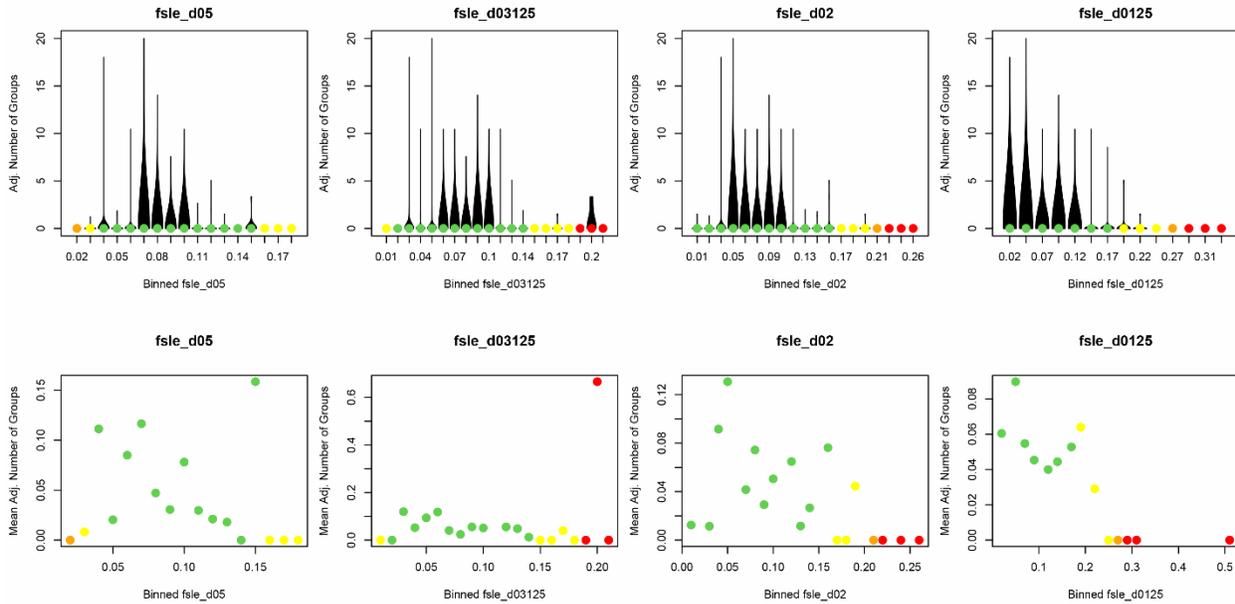


Figure 20. Relationship between variables related to finite size Lyapunov exponents and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group 10: Residual temperature – daily product

The patterns evident for the residual temperature group included highest variability in the response in the central bins, few observations in the outer bins (Figure 21). No clear pattern was evident from the mean response except for that mean responses in the lowest and highest bins were generally zero.

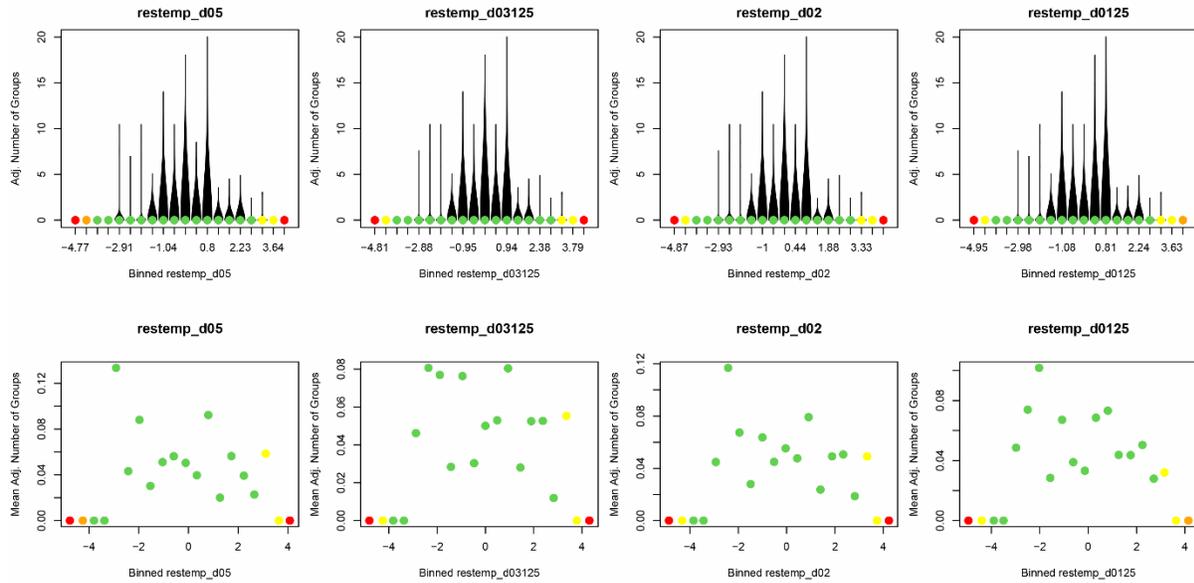


Figure 21. Relationship between variables related to daily residual temperature and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Group 11: Spatial gradient for sea-surface temperature and salinity – daily

Both covariates in this group, however, in particular *gradsst\_d025*, had the fewest observations in the highest bins (Figure 22). Both covariates had slightly increased mean responses with increasing covariate values. However, both also had the highest mean response per bin in a bin with few values.

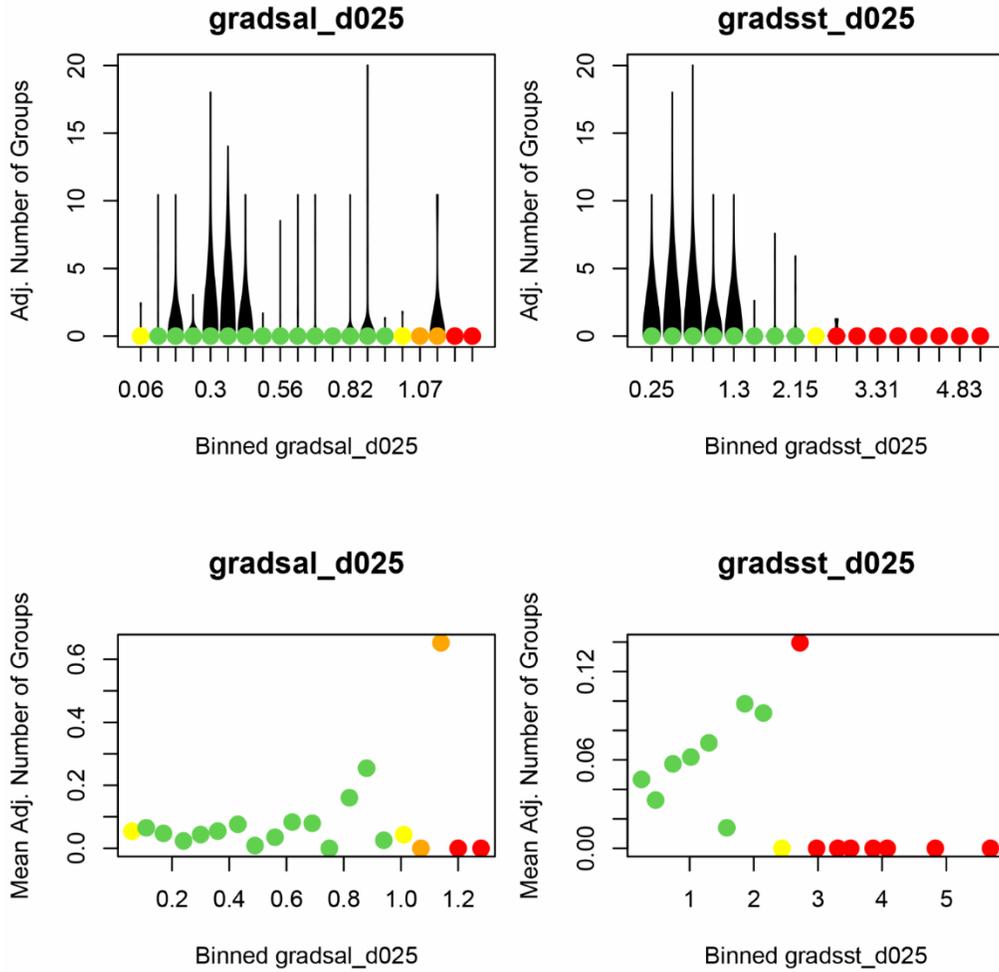


Figure 22. Relationship between variables daily spatial gradient for sea-surface temperature and salinity and the response (adjusted number of groups per segment). Top row: violin plots with median shown as coloured dots indicating the sample size within the bin (green: >100, yellow: <101, amber: <21, red: <11), interquartile ranges would show in white but all equal zero. Bottom row: means shown as coloured dots.

### Model-based analysis

The 514 transect lines were divided into 8146 segments most of which were approximately 10 km long (Figure 23). The response variable for the model (the number of groups  $N_i$  adjusted for imperfect detection) ranged between 0 and 20.01 per segment.

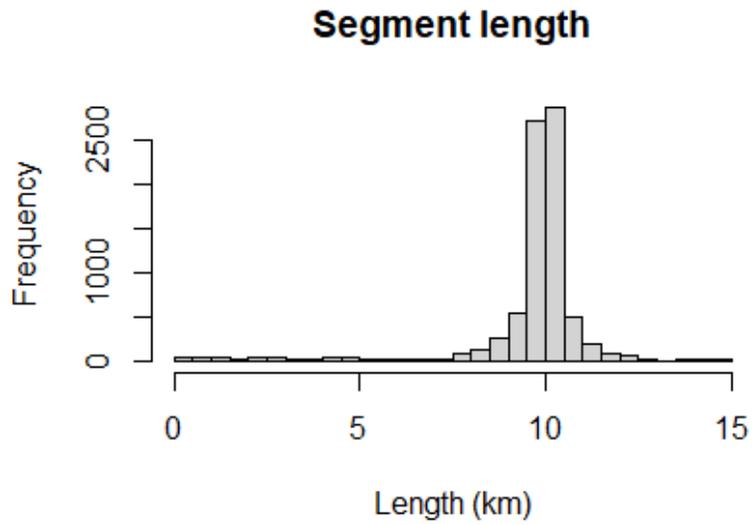


Figure 23. Distribution of segment length.

### Group Number Model

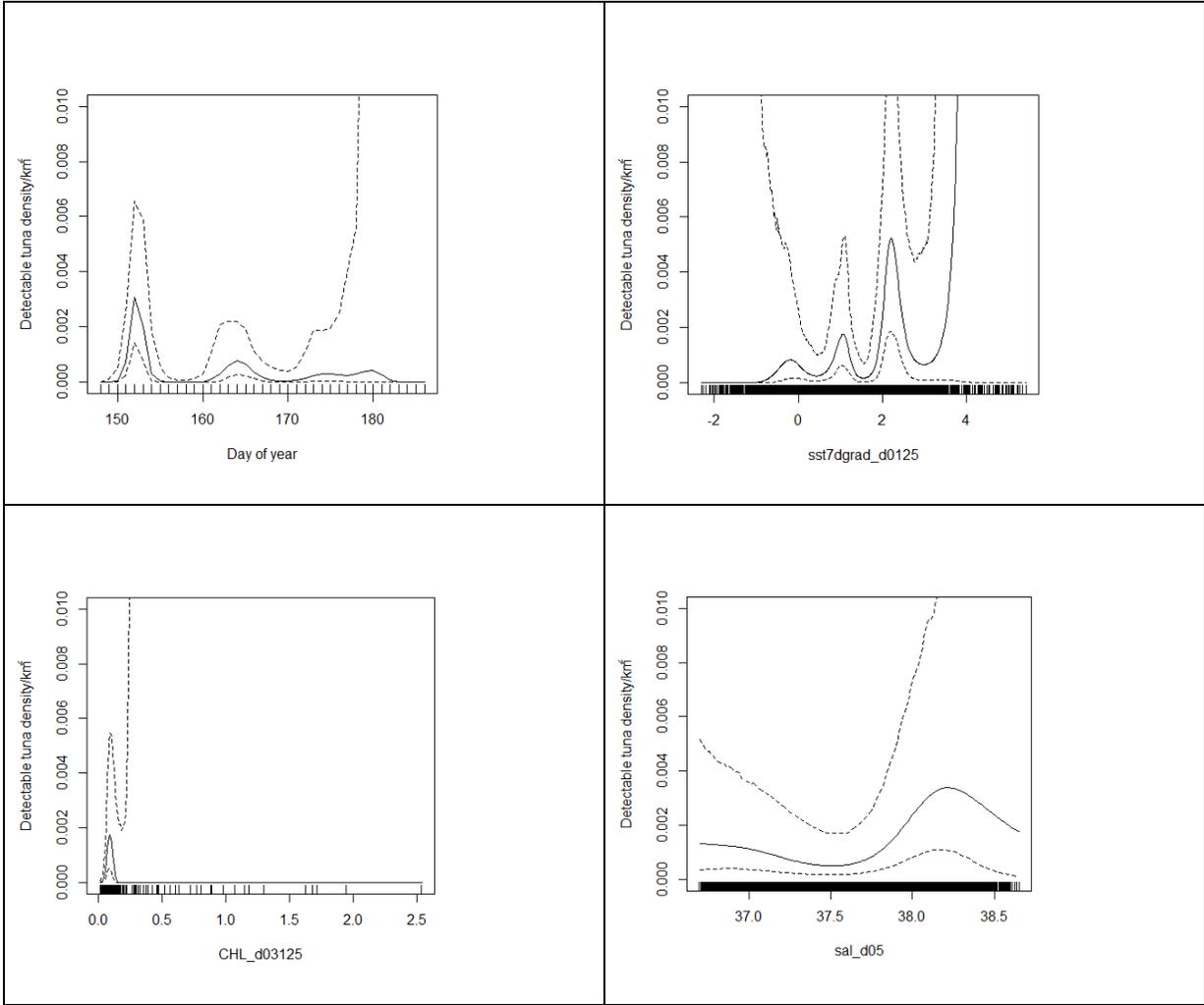
Variables and interactions for each spatial model after elimination of concurred or colinear variables are given in Table 11. Final models after model selection are also given in Table 11. The total deviance by the group number model was 30.3%. An equivalent model with the addition of a smooth of Longitude and Latitude would explain 37.4% of the deviance using an additional 33df.

Table 12. Initial variables considered and final models.

Model (dependent variable)	Initial variables considered (after vif consideration)	Final Model
<p><b>Group number model</b></p> <p><b>(Number of groups)</b></p>	<p>Slope (as a factor)</p> <p><b>1D smooths</b></p> <p>slope slope02, slope04, slope06, CHL_d03125, sst_d0125, sst7dgrad_d05, sst7dgrad_d0125, sst15dgrad_d05, sst15dgrad_d0125, mld_d02, sal_d05, sal_d0125, bathymetry, gradsal_d025, gradsst_d025 &amp; Julian day</p> <p>Smooths per slope factor were considered for bathymetry Smooths per block were considered by CHL_d03125</p> <p><b>2D Smooths</b></p> <p>Smooths of slope and bathymetry. Smooths of temperature and Julian day</p>	<p>Slope as a factor</p> <p><b>1D smooths</b></p> <p>st15dgrad_d05, CHL_d03125, sal_d05</p> <p><b>2D smooth</b></p> <p>Julian day and ,sst7dgrad_d0125</p>
<p><b>Group size model</b></p> <p><b>(Group size)</b></p>	<p>Slope (as a factor)</p> <p><b>1Db smooths</b></p> <p>Distance, slope slope02, slope04, slope06, CHL_d03125, sst_d0125, sst7dgrad_d05, sst7dgrad_d0125, sst15dgrad_d05, sst15dgrad_d0125, mld_d02, sal_d05, sal_d0125, restemp_d05, bathymetry gradsal_d025, gradsst_d025 &amp; Julian day</p> <p>Smooths per slope factor were considered for bathymetry Smooths per block were considered by CHL_d03125</p> <p><b>2D Smooths</b></p> <p>Smooths of slope and bathymetry                      Smooths of temperature and Julian day</p>	<p>distance, slope06</p>
<p><b>Biomass model</b></p> <p><b>(Density)</b></p>	<p>Slope (as a factor)</p> <p><b>1D smooths</b></p> <p>slope slope02, slope04, slope06, CHL_d03125, sst_d0125, sst7dgrad_d05, sst7dgrad_d0125, sst15dgrad_d05, sst15dgrad_d0125, mld_d02, sal_d05, sal_d0125, restemp_d05, bathymetry, gradsal_d025, gradsst_d025, Julian day</p> <p>Smooths per slope factor were considered for bathymetry Smooths per block were considered by CHL_d03125</p> <p><b>2D Smooths</b></p>	<p><b>Presence/absence component</b></p> <p>sal_d05, restemp_d05,</p> <p><b>Poisson component</b></p> <p>Slope as a factor</p> <p><b>1D smooths,</b></p> <p>sst7dgrad_d0125 , restemp_d05</p> <p>CHL_d03125 by Block</p> <p><b>2D smooths</b></p>

	Smooths of slope and bathymetry. Smooths of temperature and Julian day	Bathymetry and slope04
--	--	------------------------

Figure 24 & Figure 25 show the effect on group number on the response scale assuming mean values for the other variables not shown. There was great uncertainty in the tuna response to the more extreme covariate variables.



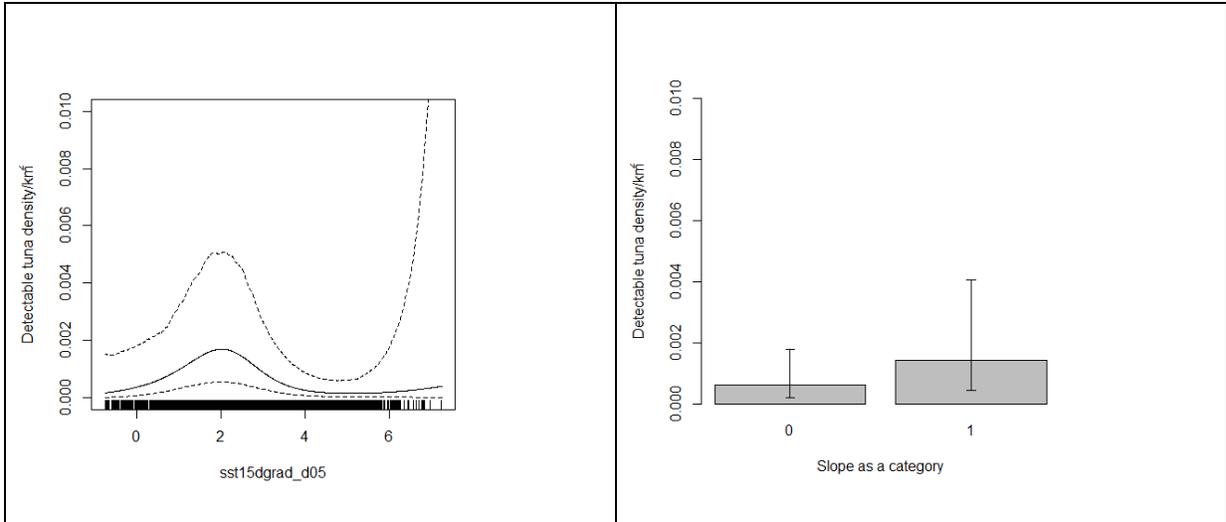
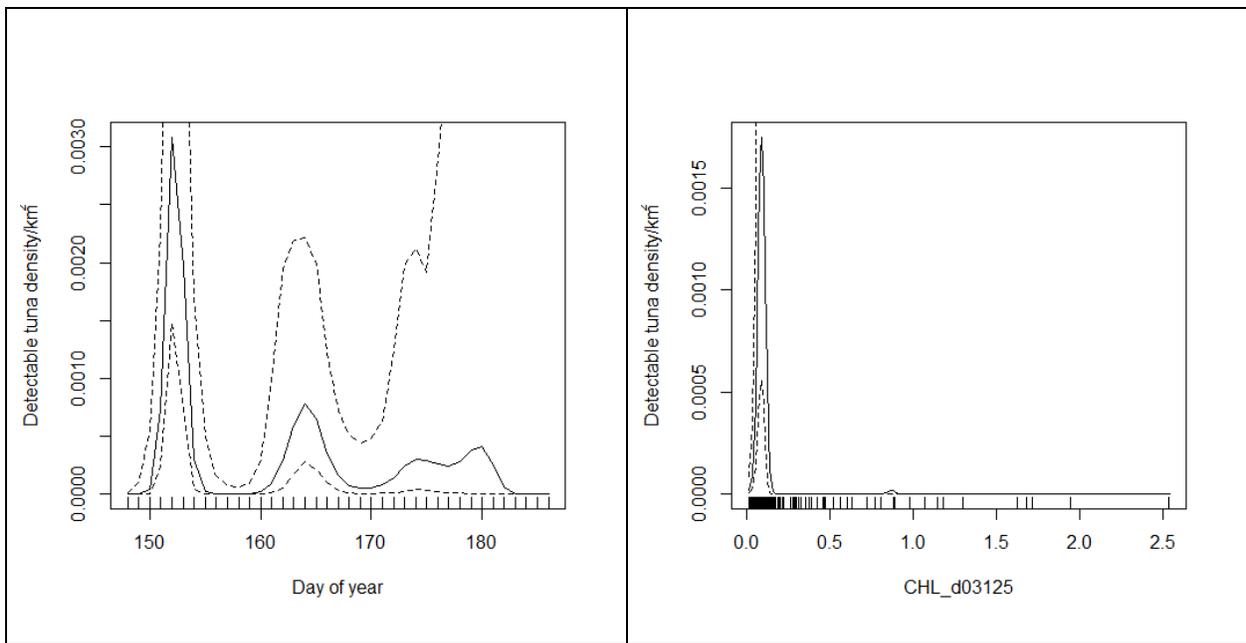


Figure 24. Effect of the different model variables, assuming a mean value for the other variables. Not that because of the presence of a 2D smooth of Day of Year and sst7dgrad\_d0125, a single dimension graph is a simplification. The rug plot indicates the distribution of the variable under consideration.



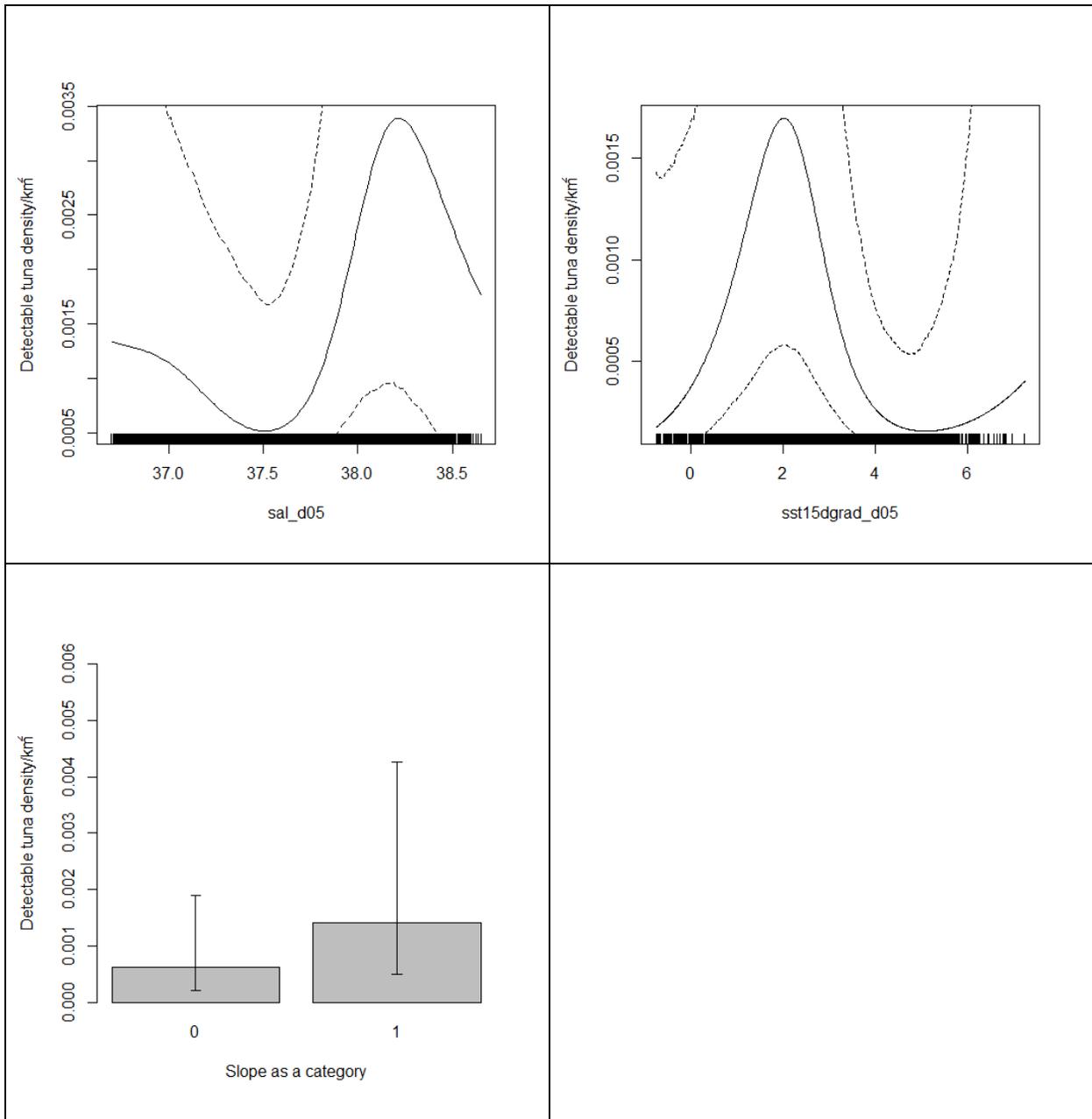


Figure 25. Effect of the different model variables, assuming a mean value for the other variables, zooming in on some of the variables in Figure 24 so not to scale. Note that because of the presence of a 2D smooth of Day of Year and sst7dgrad\_d0125, a single dimension graph is a simplification. The rug plot indicates the distribution of the variable under consideration.

### Best model for group size

After eliminating collinear covariates and model terms with non-significant p-values, we arrived at a model consisting of slope06 and distance as linear (on the scale of the link function) terms (Table 12). The total deviance explained was low: 4.0%.

The effect of each variable on group size is given in Figure 26. Note that when considering the total range of slope 06 over the survey, rather than just the sightings range there is considerable uncertainty.

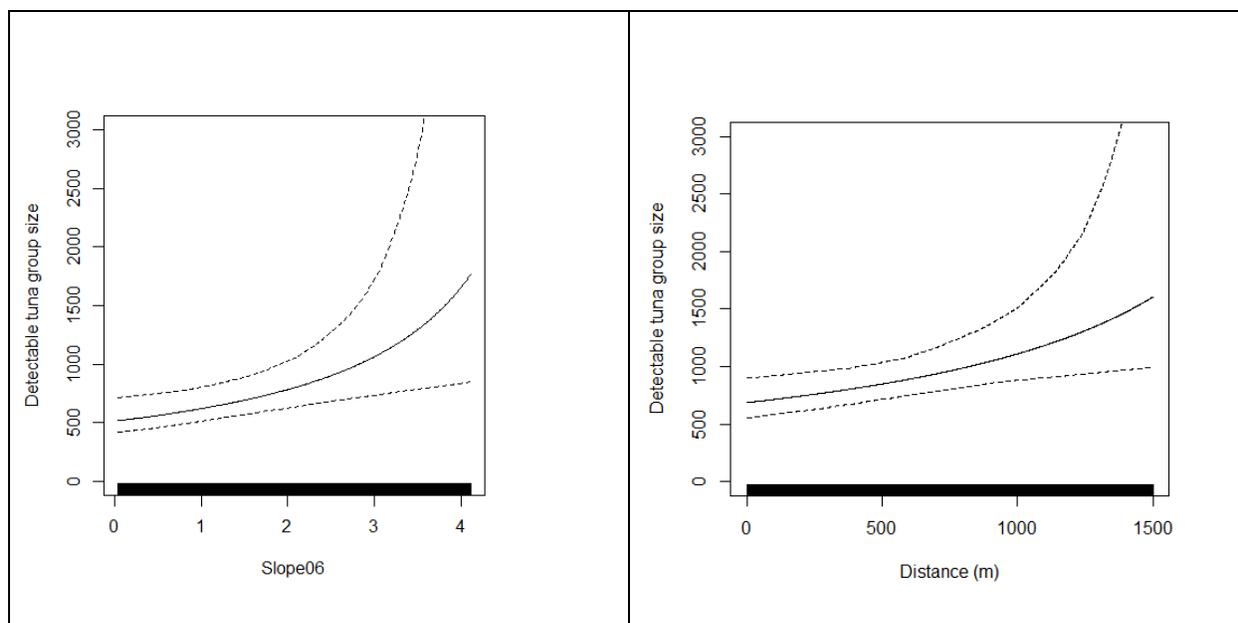


Figure 26. Effect of the different, group size model variables, assuming a value of zero in the case of slope06 and the mean of slope 06 for distance. The rug plot indicates the distribution of the variable under consideration.

### Estimating fish numbers and density

The group number model (Figure 24) suggested that the optimum date for tuna group numbers was c. May 31<sup>st</sup> (Julian day 151) so estimates were made for this date using the predicting value of the covariates on this date. These results were multiplied by the predicted values from the group size model (which were time invariant) to produce the estimates given in Table 13 and Figures 27 – 31, with the time series in Figures 33 – 34. Ninety-five percent confidence intervals were obtained from the bootstrap.

Tuna estimates for Block C, were inflated by extreme values of the predictors in the south-west corner (Figures 27 to 31, yellow region) caused by the great uncertainty in regions of very high chlorophyll (Figure 34). Alternative estimates were generated by omitting this region (bracketed value in Table 13, Figure 33).

Tuna distribution is generally patchy, but block A typically had more tunas than the other blocks. There is no clear and consistent areas of high abundance per block in between years (Figures 27-31).

Table 13. Estimated number of individuals  $N$  per Block and 95% CIs. The bracketed UCLs for Block E reflect omission of the south-west locations with unrealistic predictions. CVs are for this smaller region for Block E for day of year 151.

Block	Year	Abundance	CV	LCL	UCL
A	2017	158700	0.51	96800	345500
	2018	176700	0.53	113000	355200
	2019	274800	0.29	191800	449700
	2021	437700	0.32	280200	788200
	2022	221500	1.51	66800	1106800
C	2017	53600	0.51	36000	126700
	2018	131300	0.72	71700	392600
	2019	355000	1.82	94500	2087500
	2021	158800	0.41	117900	329600
	2022	52900	0.84	30500	132900
E	2017	174800	(1.47)	72600 (51000)	Inf (873100)
	2018	101100	(0.33)	69900 (69000)	$2.594 \times 10^{35}$ (177900)
	2019	355500	(1.67)	91000 (81700)	$1.589 \times 10^{84}$ (2028400 )
	2021	111200	(0.53)	64800 (61200)	$1.574 \times 10^{58}$ (246100)
	2022	42400	(0.42)	30100 (27900)	Inf (82200)

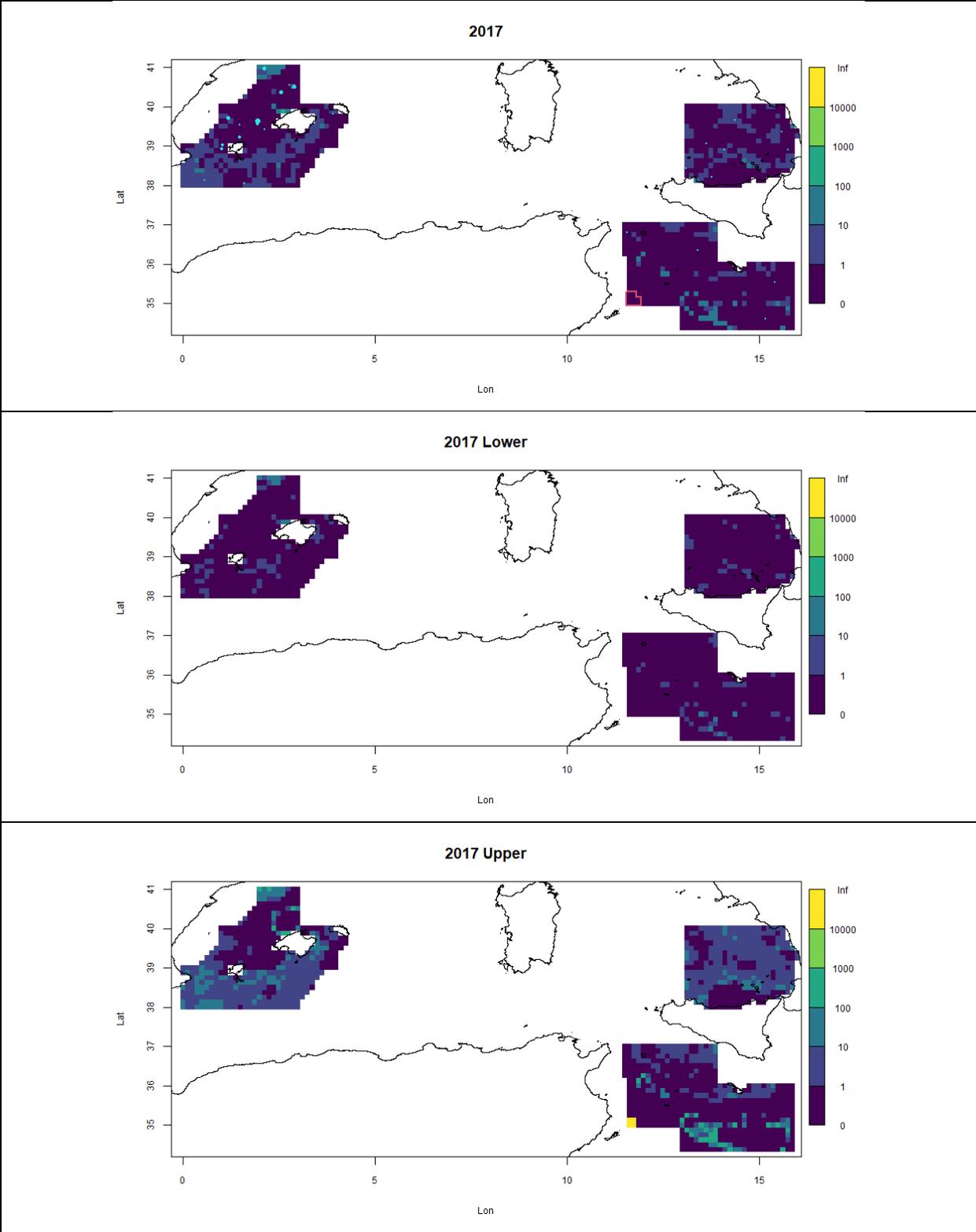


Figure 27. Surface tuna density estimates for day 151, 2017 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

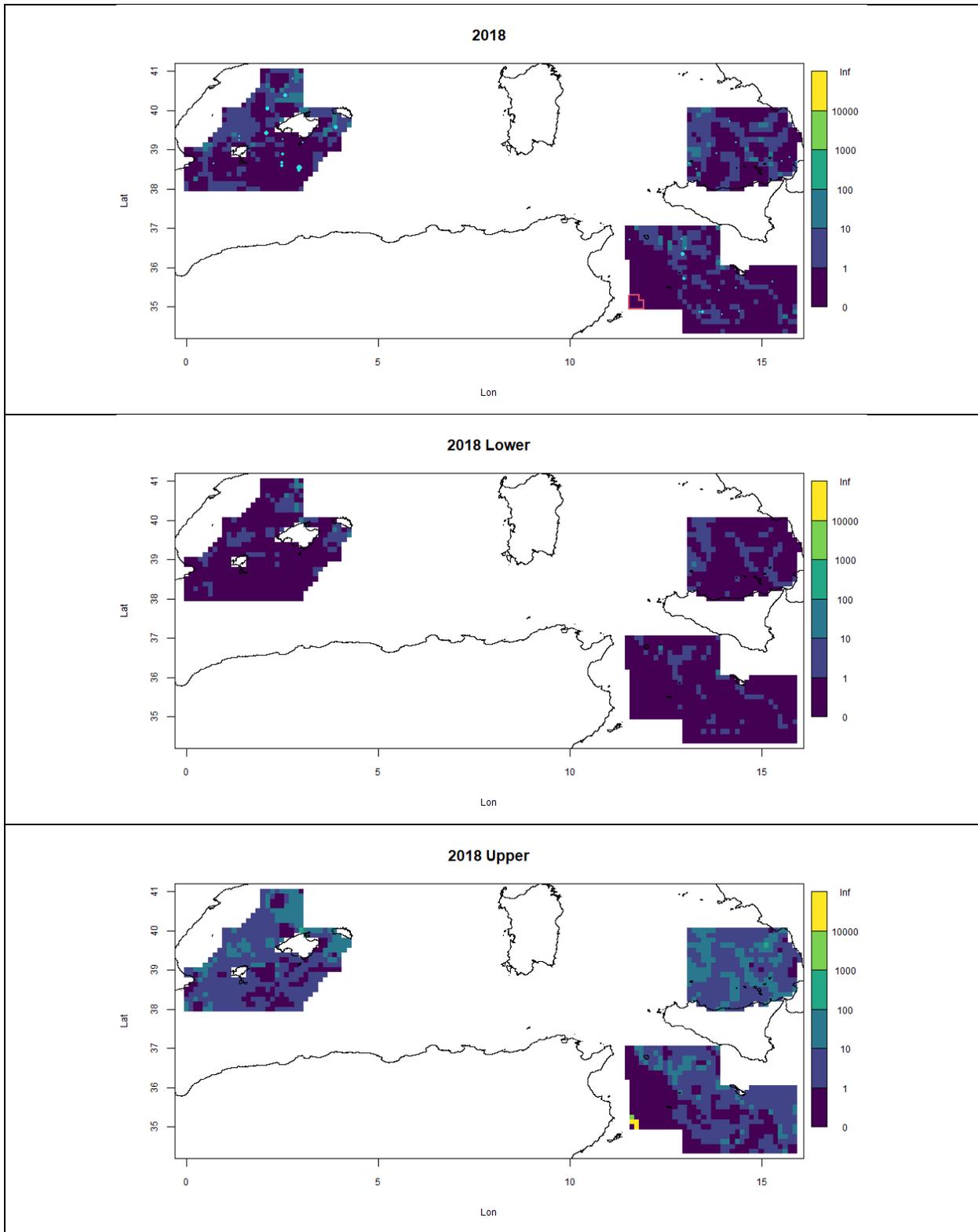


Figure 28. Surface tuna density estimates for day 151, 2018 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

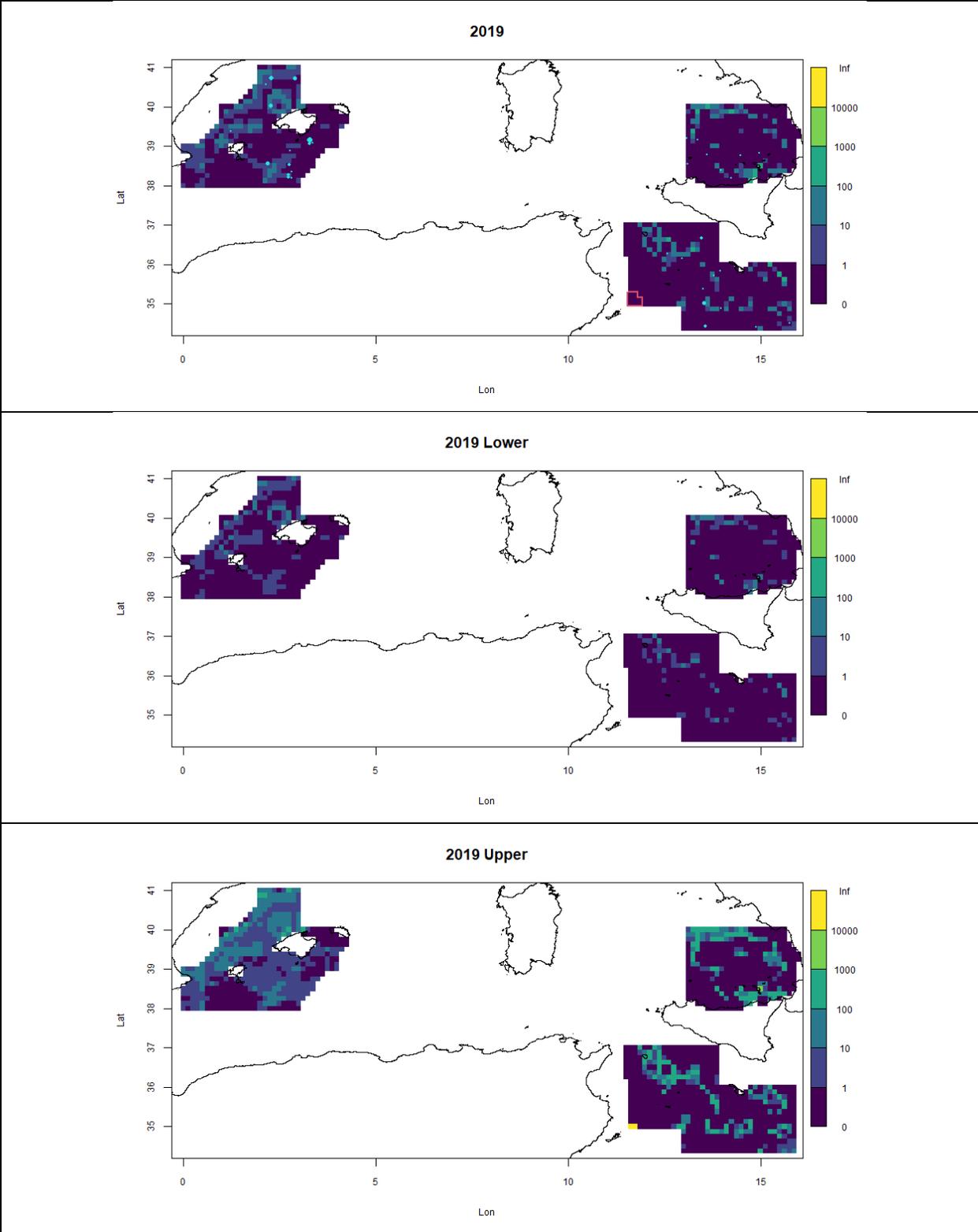


Figure 29. Surface tuna density estimates for day 151, 2019 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

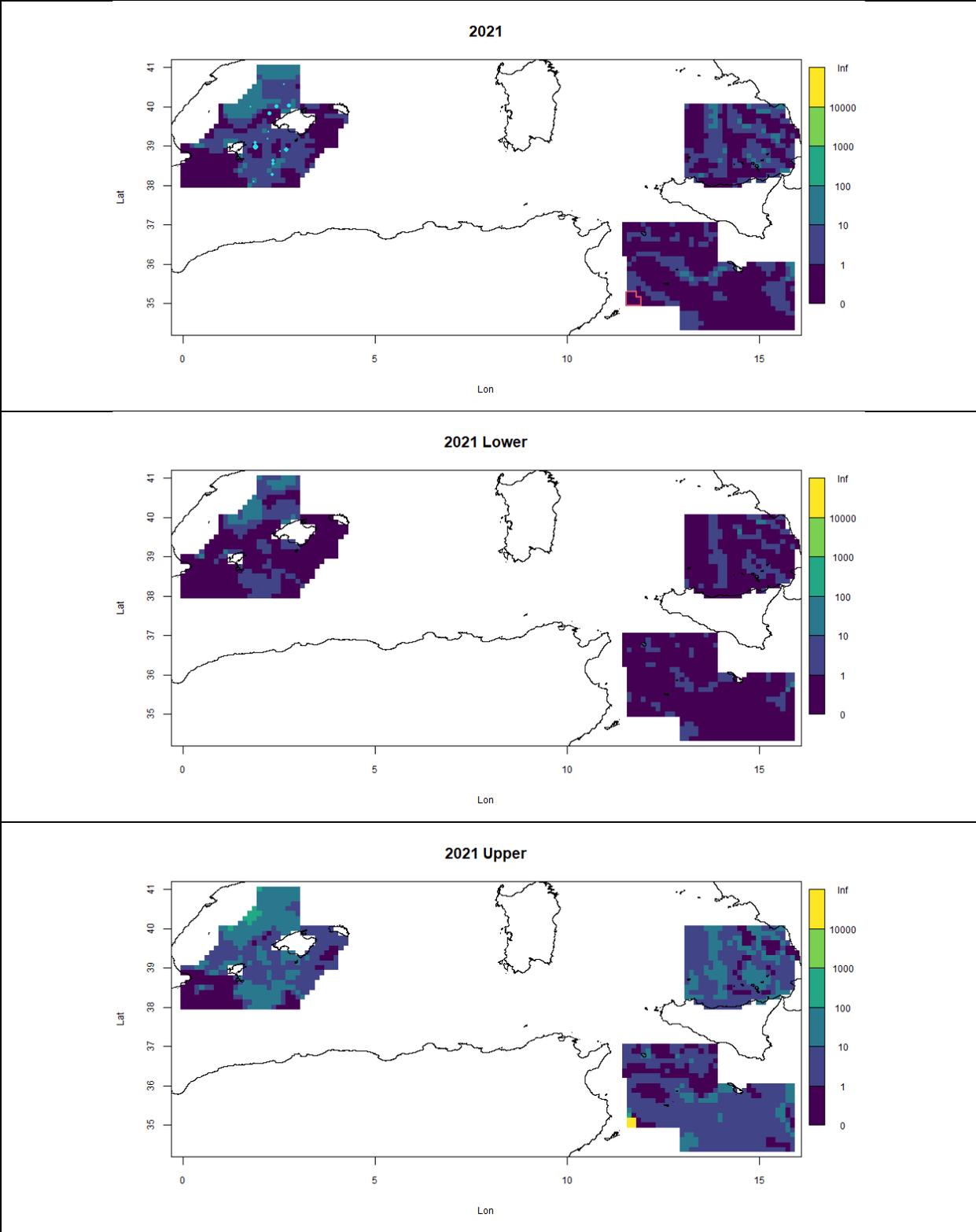


Figure 30. Surface tuna density estimates for day 151, 2021 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

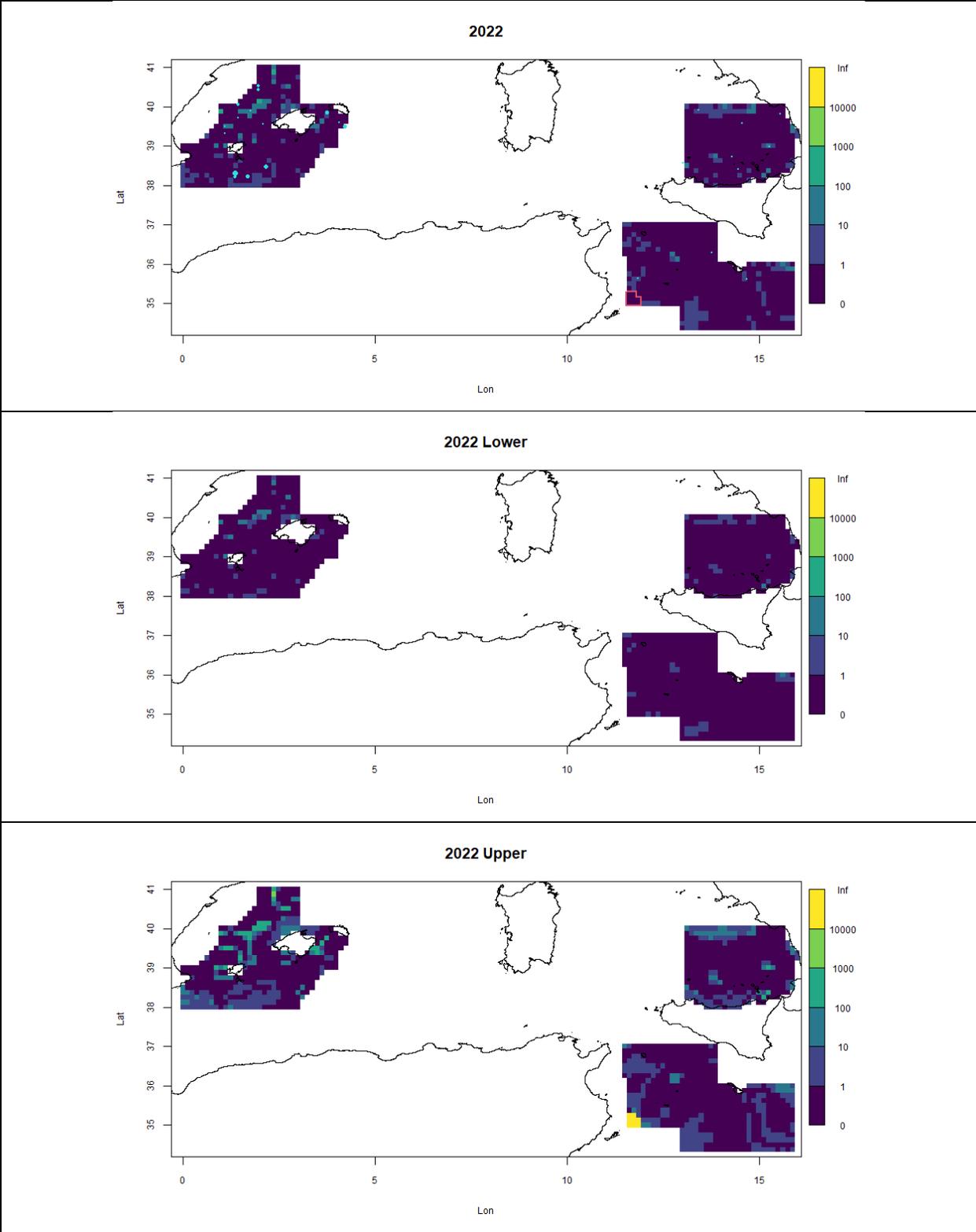


Figure 31. Surface tuna density estimates for day 151, 2022 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

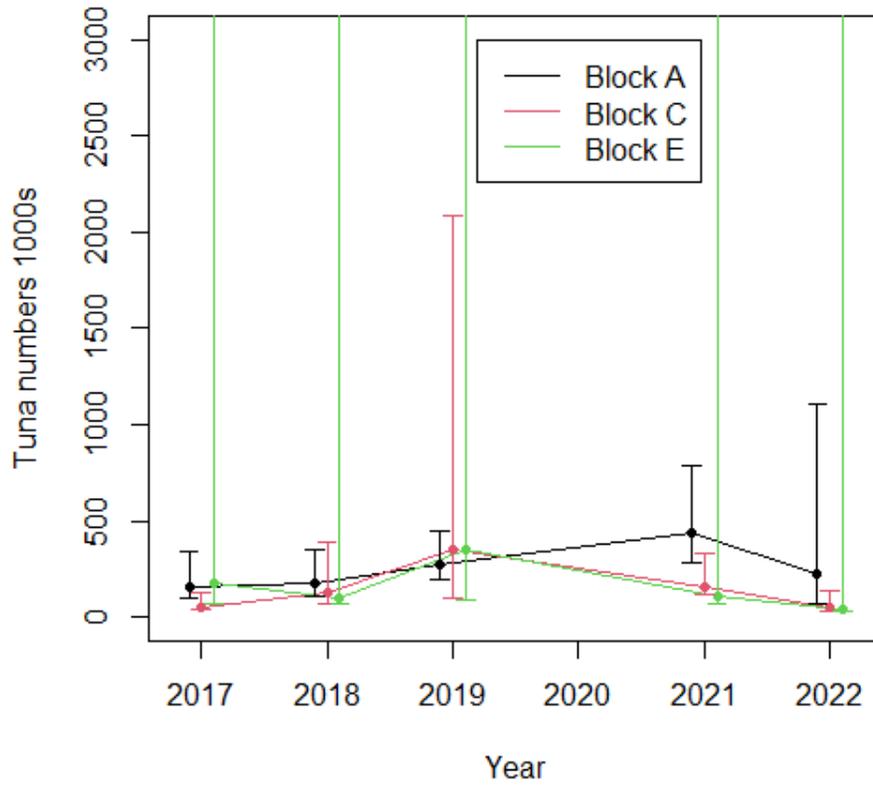


Figure 32. Time series of tuna density estimates for day 151, 2017 – 2022. Note there was no surveying in 2020. High upper bounds for Block E are caused by the predictions in the south-west region and region associated with a high chlorophyll count.

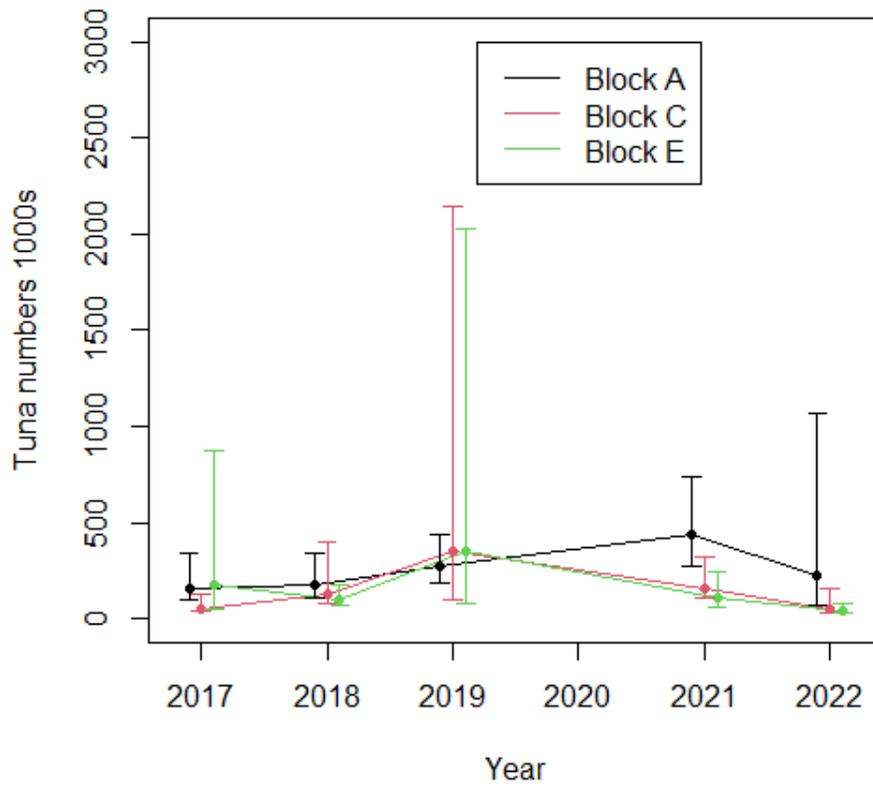


Figure 33. Time series of tuna density estimates for day 151, 2017 – 2022. Note there was no surveying in 2020. These limits omit the highly uncertain south-west region of Block E.

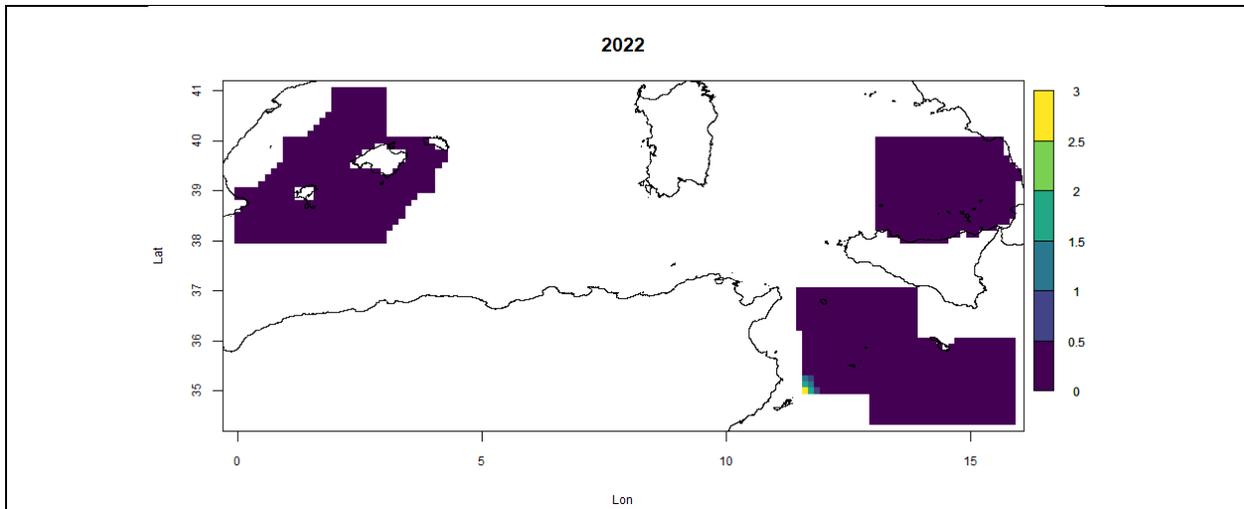


Figure 34. Spatial maps of estimated chlorophyll d30125 ( $\text{mg m}^{-3}$ ) for day 151, 2022. Note the “hotspot” to the south-west of Block E.

Estimates were also undertaken for later on in the sampling season, to see if there was a large-scale shift in spatio-temporal abundance. The figures for day 163 (~12 June) are given in Table 14 and illustrated in Figure 40. Spatial changes in density estimates are given in Figures 35-39.

Here uncertainty was very much higher, caused not by chlorophyll so much as by uncertainty in the response to the dynamic temperature variables. The distribution of hotspots was patchy and there was no consistent location of hotspots between these two prediction dates except for the SE area E hotspot. Estimated abundance in Block A was higher than for day 151, especially in the recent years (Figure 33 vs Figure 40).

Table 14. Estimated number of individuals per Block and 95% CIs. The bracketed UCLs for Block E reflect omission of the south-west locations with unrealistic predictions. CVs are for this smaller region for Block E. For day of year 163.

Block	Year	Abundance	CV	LCL	UCL
A	2017	642300	0.23	454800	998200
	2018	185400	0.56	99800	489600
	2019	18100	Large	7100	$2.2 \times 10^7$
	2021	2311400	Large	346300	$2.2 \times 10^{10}$
	2022	169200	0.30	110200	305900
C	2017	20000	0.76	9300	63700
	2018	45200	Large	23300	571500
	2019	95300	1.35	36200	455700
	2021	75500	0.39	44800	159300
	2022	700	Large	100	$5.2 \times 10^{10}$
E	2017	34400	Large	26600	$7.4 \times 10^{182}$
	2018	18800	$\infty$	12900	$\infty$
	2019	120500	$\infty$	95700	$\infty$

	2021	20400	$\infty$	16000	$1.0 \times 10^{296}$
	2022	0	*	0	$\infty$

\*Not defined

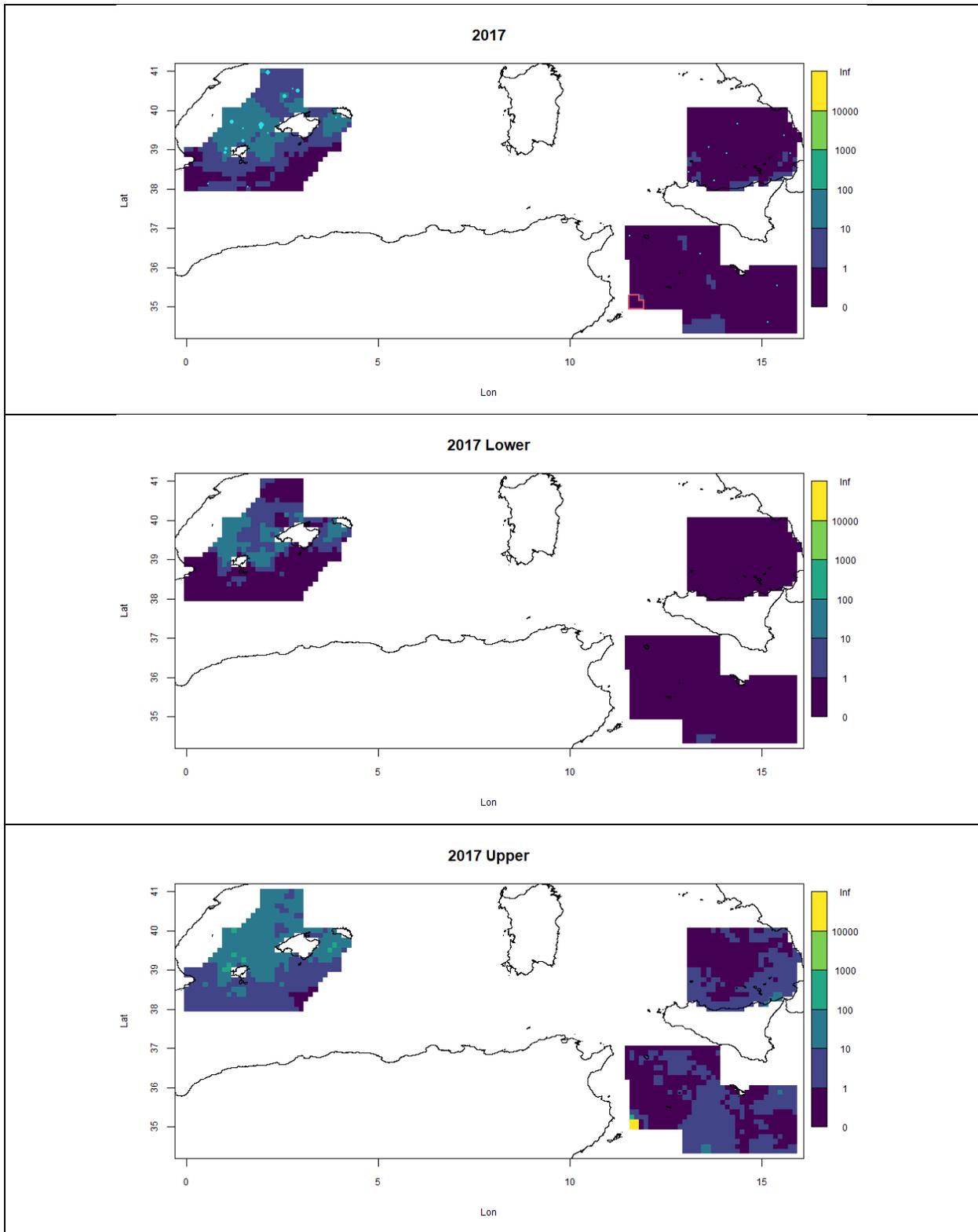


Figure 35. Surface tuna density estimates for day 163, 2017 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

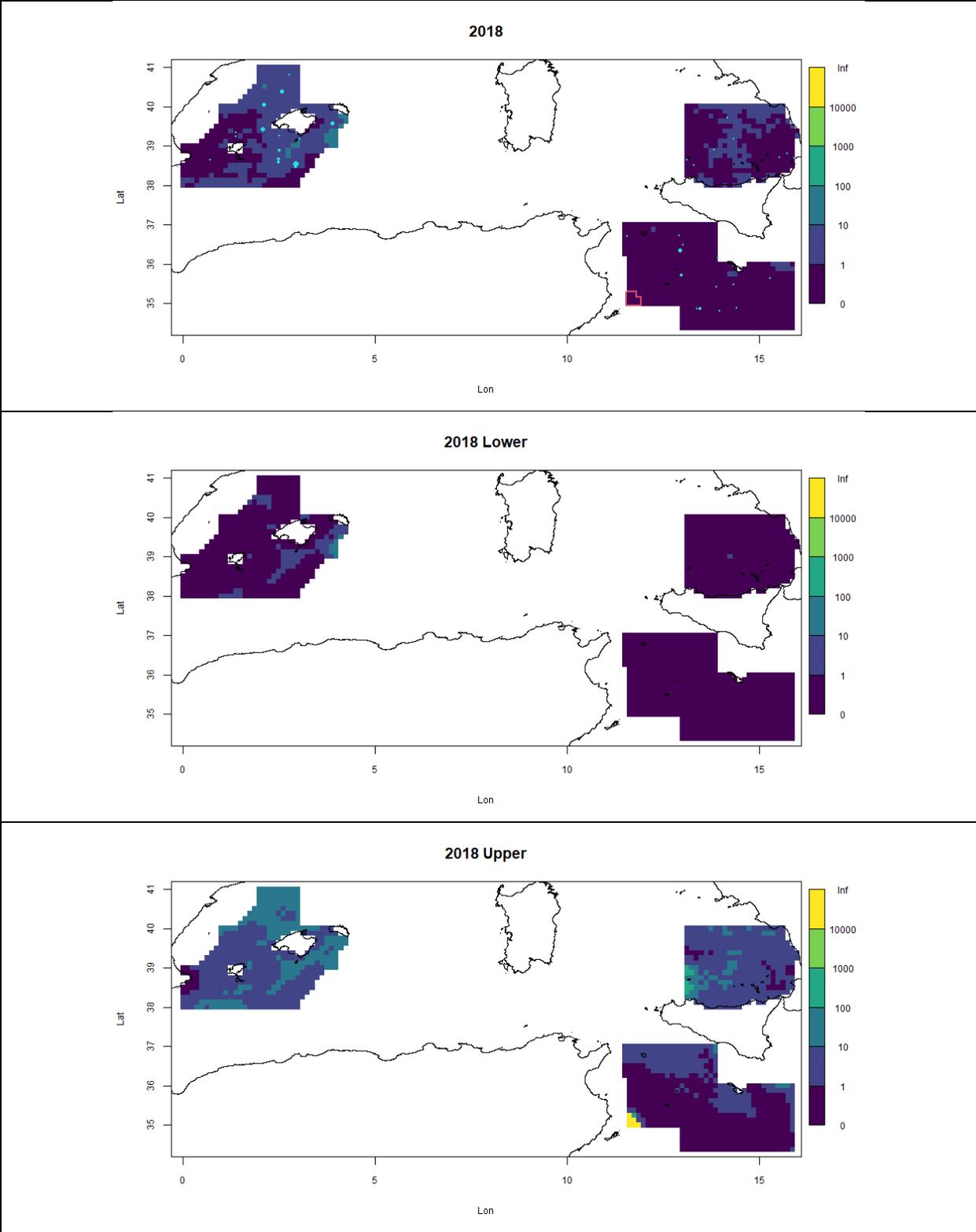


Figure 36. Surface tuna density estimates for day 163, 2018 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

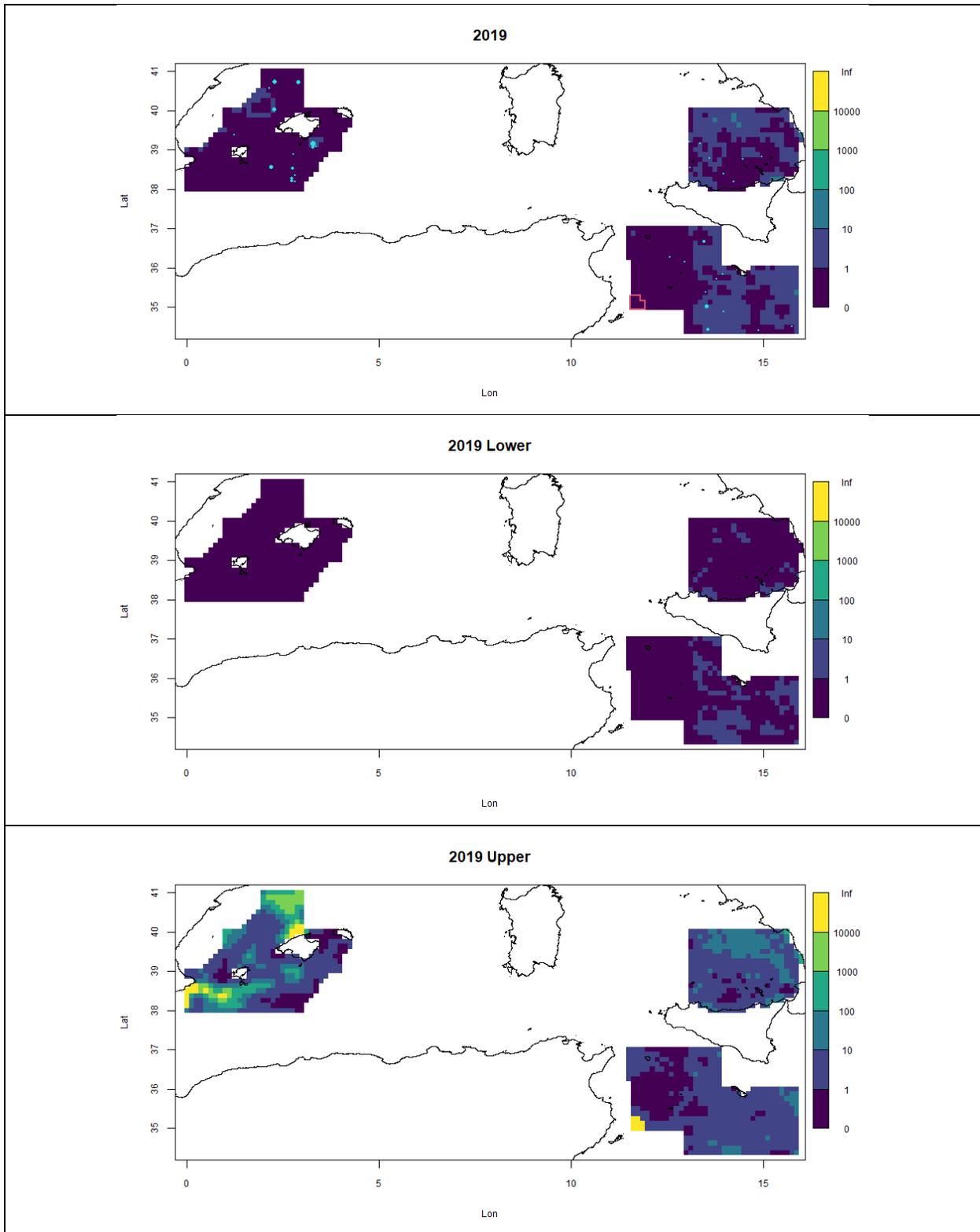


Figure 37. Surface tuna density estimates for day 163, 2019 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

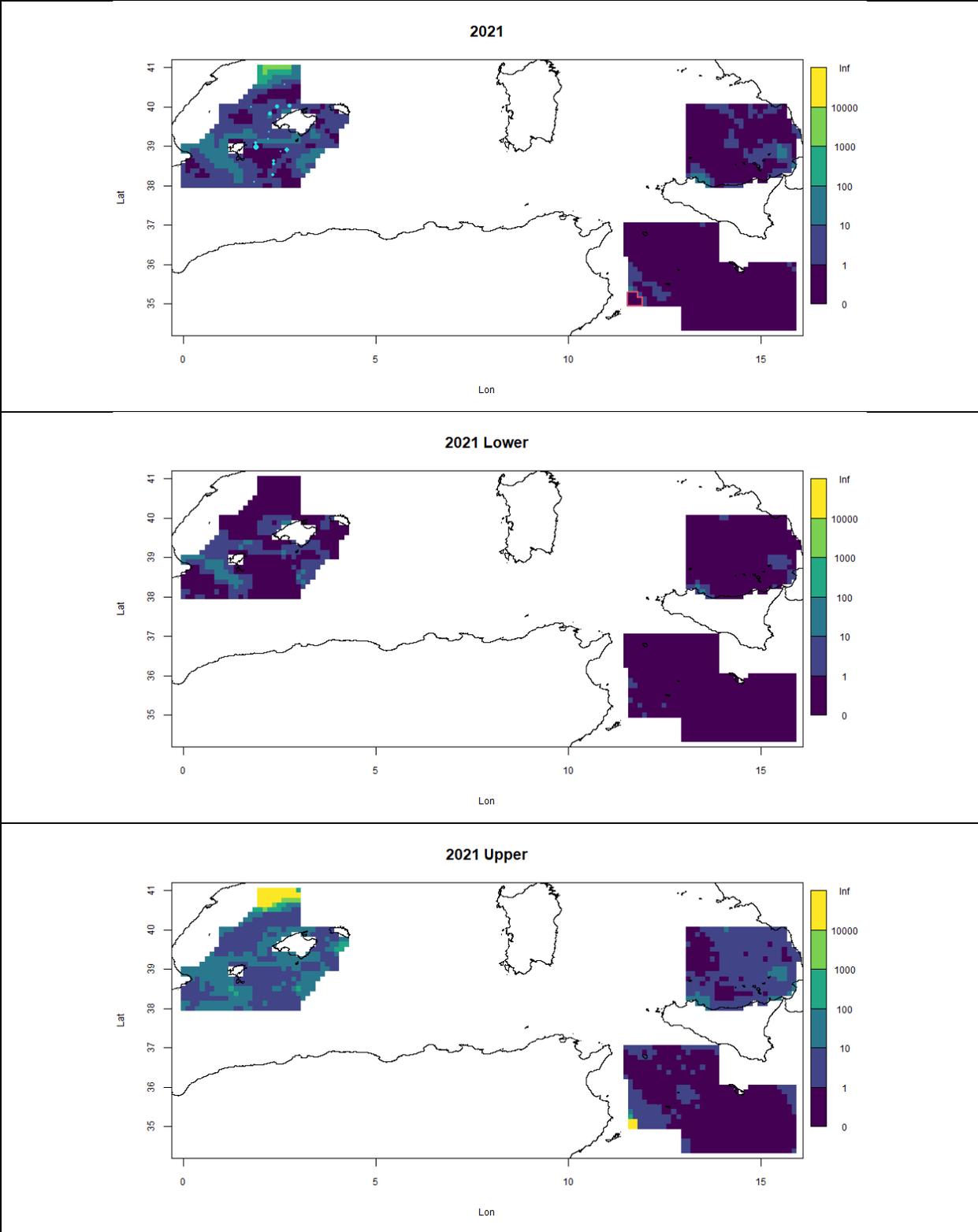


Figure 38. Surface tuna density estimates for day 163, 2021 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

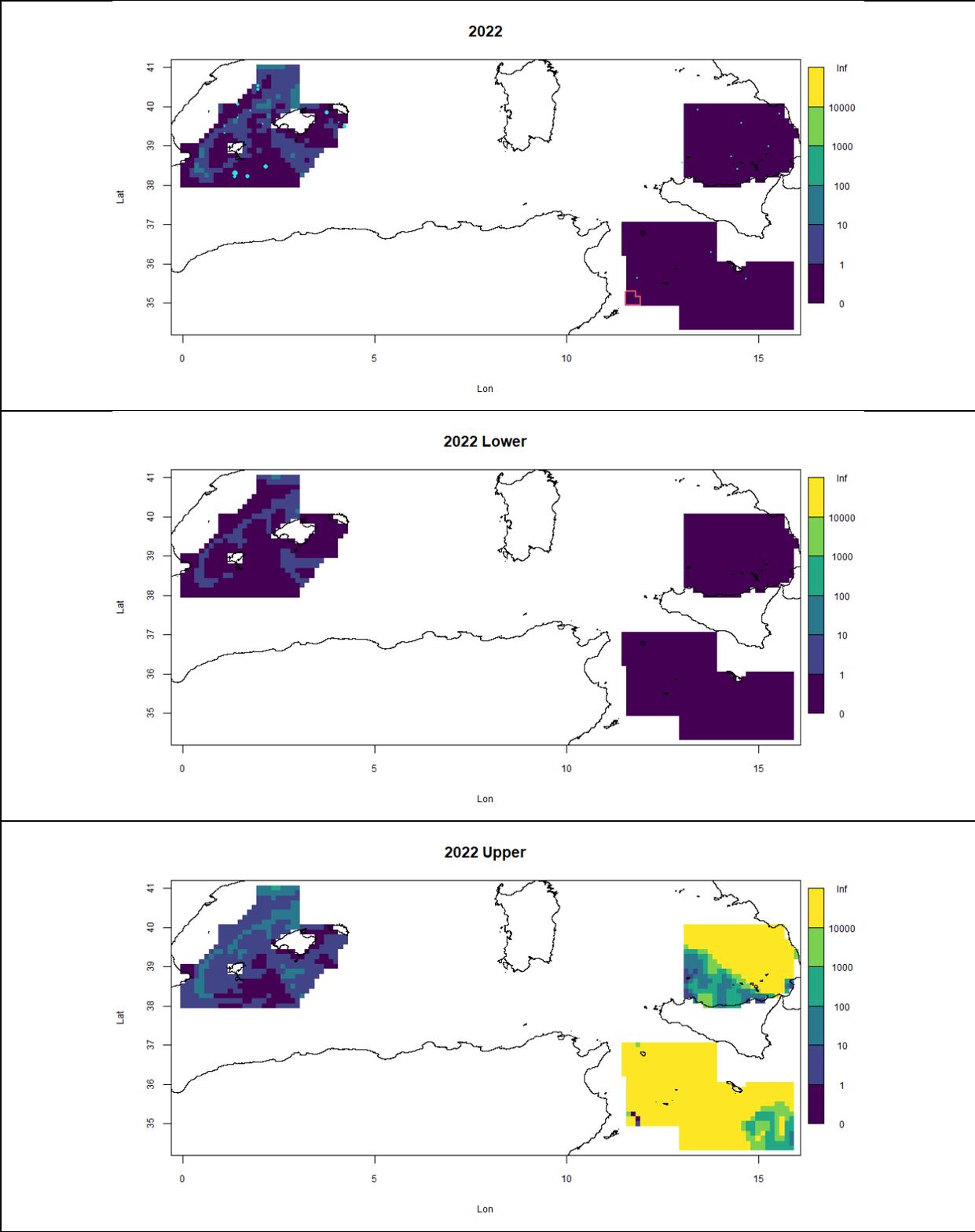


Figure 39. Surface tuna density estimates for day 163, 2022 (top) with associated lower (middle) and upper (bottom) confidence bounds. Turquoise dots indicate seen tunas (area proportional to detection adjusted group numbers). The red outline illustrates the highly uncertain region omitted for some of the estimates.

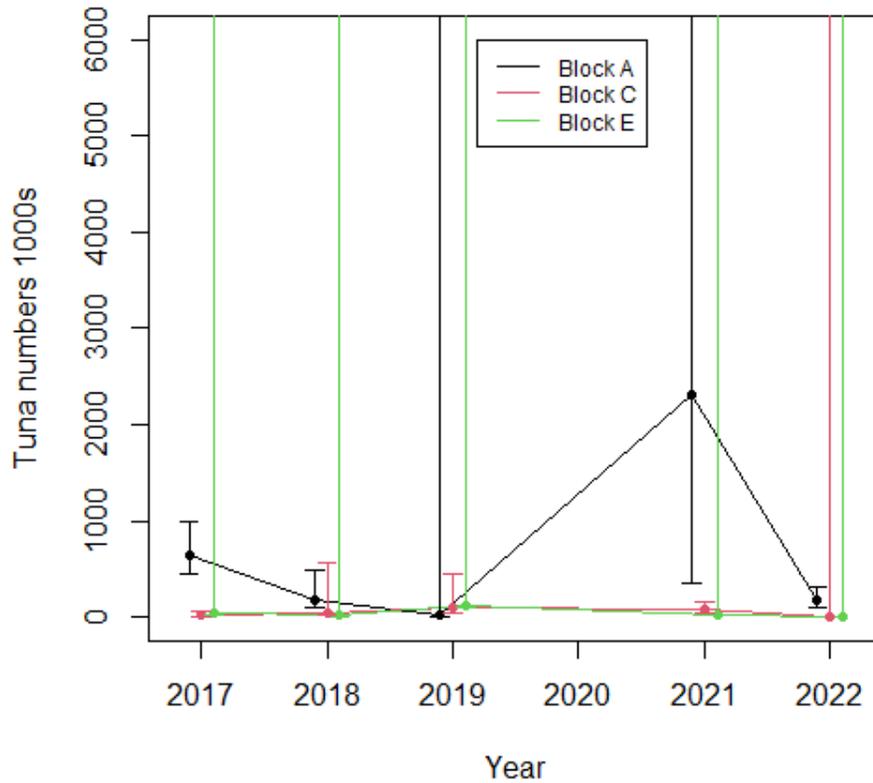


Figure 40. Time series of tuna density estimates for day 163, 2017 – 2022. Note there was no surveying in 2020. These limits omit the highly uncertain south-west region of Block E.

Figure 41 shows comparison of BFT abundance estimates between design-based and model-based approach for the two prediction days (151 and 163). The estimates from model-based approach are higher than design-based approach for all years and blocks but highest discrepancies are for block A.

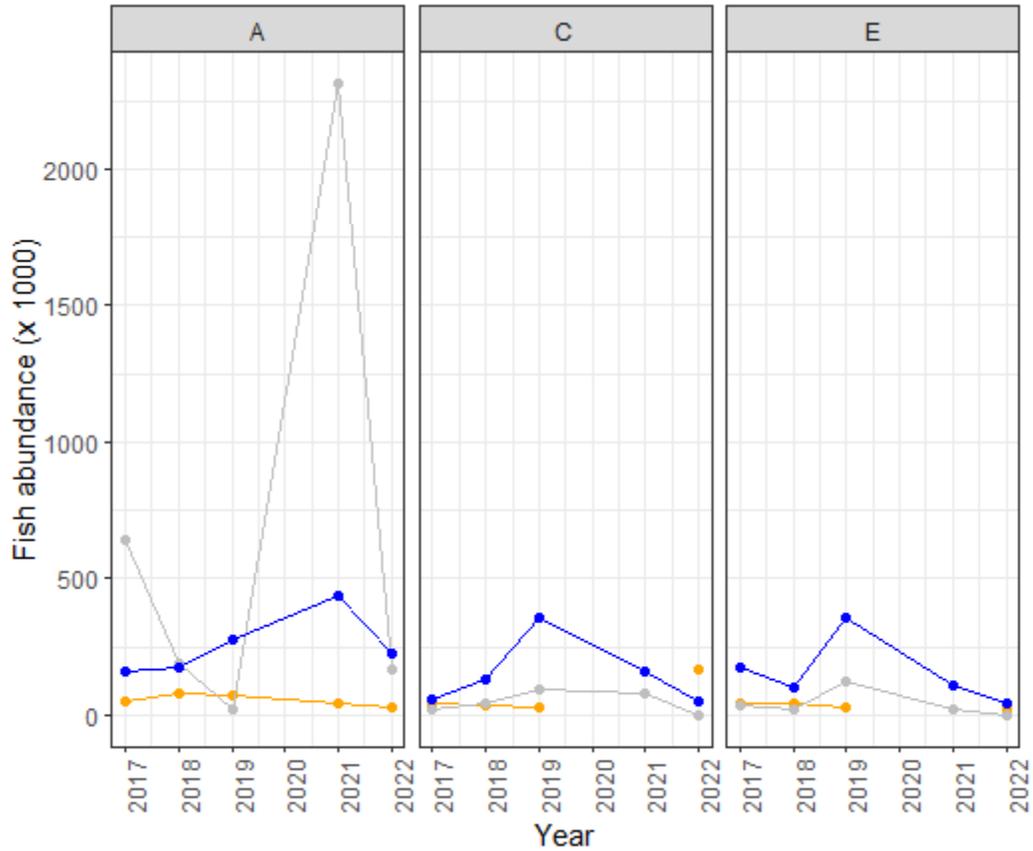


Figure 41. Comparison of mean estimates of BFT abundance for each block and year between design-based estimates (orange), model-based estimates for day 151 (grey) and model-based estimates for day 163 (blue).

Figure 41 shows a comparison of the design and model-based estimates for each block and year. Model based point estimates are generally much higher than their design based equivalents but there is much greater uncertainty (shown above) caused by the putative hotspots caused by undersampled covariate values.

## Discussion

It should be stressed that the figures given here are based on estimates from fish observed at the surface and so may be very different to the actual number/biomass of fish. Further behavioural changes in the fish in response to environmental conditions may cause changes in estimated numbers which are not proportional to actual changes in abundance.

## Design- and model-based approaches

Including *company* as a covariate in the detection function was caused by varying distributions of observed distances between companies (Figure 42).

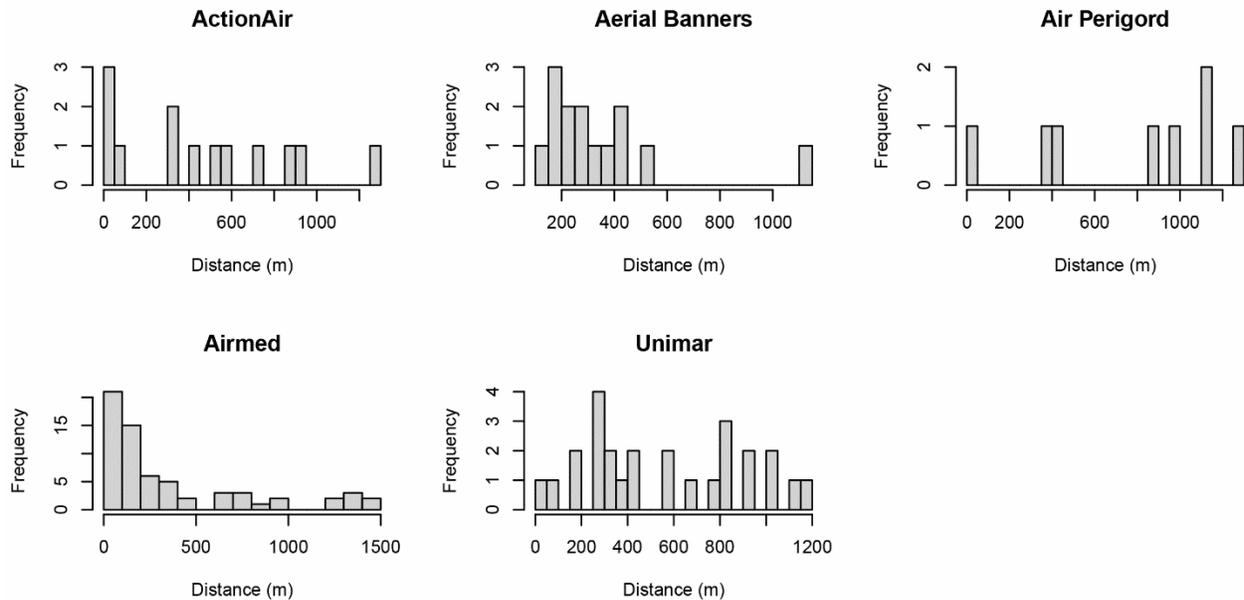


Figure 42. Histograms of observed distances for each company.

The overall shape of the average detection function was determined by company Airmed who made the majority of detections during the surveys included in this study (Table 15). The final selection of the detection function with *company* and the *log-size* as covariates was driven by this company also making most of the detections with small school sizes. Out of 31 detections with schools of less than 100 individuals, 29 were made by company Airmed. This was also the only company that spotted any schools of 10 or less individuals. This was not surprising as this company also made over 50% of all detections, despite only conducting 30% of the effort (Table 15). Including company or airplane type in the detection function is consistent with detection functions used in the analysis of aerial surveys from previous years (Chudzinska et al. 2021, Chudzinska et al. 2022).

Table 15. Proportion of sightings and effort shown by the five companies conducting *the surveys*.

	ActionAir	AerialBanners	AirMed	AirPerigord	Unimar
Sightings	0.10	0.11	0.52	0.06	0.21
Effort	0.19	0.15	0.30	0.07	0.30

The inclusion of company as well as airplane type in estimated detection function in the previous years (Chudzinska et al. 2021, Chudzinska et al. 2022), may indicate that more standardized survey protocol should be implemented. Whereas the way of estimating distance to spotted groups has been standardized in surveys from 2017 onwards (Figure 5), the estimation of group sizes and biomass is not standardized. The small group sizes reported by Airmed only apply to Block A, but the two remaining companies surveying the same block

(Table 1), did not report such small groups to the same extent as Airmed. We, therefore, recommend a more consistent survey protocols across companies and blocks in terms of reported group sizes.

One advantage of the spatial regression approach as undertaken here allows estimation into regions that were not necessarily temporally or spatially surveyed. This could be critical if for example, fish abundance varies throughout the sampling period, which might lead to a design-based estimate being inaccurate. Obviously, such extrapolations can be misleading if the covariate is not well represented in the data, or the point estimates may be reasonable there can be great uncertainty. In theory, a well-supported model should lower the overall uncertainty in estimates compared to a design-based approach as it explains more of the variation.

Both the effect plots (Figure 26) and the upper CI intervals for spatial changes in density (Figures 27-31 and 35-39) show high uncertainties at the extreme values of the covariates. Predictions for such values should be interpreted with caution or can be set (or binned) to one maximum value. CVs generated from a bootstrap can be misleadingly high as just one aberrant bootstrap replicate can massively increase variance. But such a replicate would not affect the estimated confidence interval. Nevertheless, the uncertainty for some areas at particular times was high.

In the two days considered here (151 and 163), uncertainty was generated by the modelled response of tuna numbers to high levels of chlorophyll in the south-east region of Block E. On day 163 this effect persisted but for other blocks there was uncertainty in the temperature related variables. This generated high uncertainty more generally.

There are several ways to deal with this. Firstly, one could not predict in regions where little is known about the response to the covariate values. However, a more practical solution in the context of GAMs here would be employ fewer degrees of freedom in the models (or increase the penalty for wiggleness). This would reduce the general uncertainty for day 163. However, it would not have prevented to the uncertainty associated with high chlorophyll hotspot. Further work can investigate this further.

The regression approach also allows insights into the biological drivers of the distribution (see below for a discussion). Adding purely spatial covariates (e.g. Lon. And Lat.) can increase explanatory power (as was the case for the group abundance model here) at the cost of biological interpretability. Likewise including seemingly functionally similar variables (that are sufficiently distinct to supply additional information in model fit) and so increase explanatory power may reduce their overall direct interpretability as they are the effect of the variable given another very similar variable.

The highest abundance across all years and blocks was estimated for 31<sup>st</sup> May which corresponds to the beginning of surveying period or even before the start of the surveys, especially in the recent years (Table 1). This may indicate that current survey dates may occur after the peak in abundance of BFT and earlier start of the surveys is, therefore, recommended. The new survey dates may have to be designed separately for each block given higher biomass

estimates for day 163 in Block A and the unusually late survey time in that block in 2021 and 2022.

Overall these models represent a first attempt to try to produce a universal model for half of the Mediterranean. Diagnostics were not ideal (Appendix) and uncertainties high suggesting a lot of variation remains to be explained and also the tentative nature of the models at this stage. Because the objective here was to try to produce a universal (pan-Mediterranean) model, apart from one case (chlorophyll) functions were not allowed to vary by block. It is noticeable that for Block A, the results are very different from the recent universal model compared to the previous estimates which might indicate the specific conditions of blocks should be considered separately. This assumption could be more formally tested in the future.

### **Environmental drivers of BFT abundance**

The drivers of BFT abundance in the western and central Mediterranean during the spawning season and based on aerial surveys are environmental variables that have been identified in the literature as being ecologically meaningful for tuna in the reproductive season.

Bluefin tuna aggregations during the reproductive season are preferentially observed close to the shelf-slope, where the higher values of slopes are detected. This is in line with studies in the Gulf of Mexico, another prominent spawning area for BFT, where hotspots of spawning individuals have previously been found in slope waters (Block et al. 2005). This is consistent with previous studies that found that, in the Balearic Islands, BFT are preferentially found close to the surface waters, but they also perform some deep excursions (Aranda et al. 2013). The topography is known to have an influence on the regional currents and water masses, which could be driving the distribution of tuna. Three mechanistic processes could explain the presence of BFT in the areas influenced by steep slopes. These are: association with surface currents entering the spawning areas and are used by BFT during their displacements; preference of areas with colder deep waters where tuna can reduce metabolic needs by lowering down their temperature; and trophic interactions, as deep areas associated to the slopes are also used by other large migratory species for feeding. This result is especially meaningful from the output of the group-size model.

Salinity is an indicator for frontal areas. In the Balearic Islands, a front is located at the encountering of the low salinity surface Atlantic waters and the resident Mediterranean waters (more saline), at salinity values around 37.5 (Balbín et al. 2014). This is the preferable BFT spawning ground, as indicated by their preference for waters with a salinity range of 36.9 to 37.7 (Alemany et al. 2010, Reglero et al. 2012). The spatial distribution of BFT larvae has been shown to be linked to salinity in other spawning grounds, like the Gulf of Tunisia and the Gulf of Mexico (Muhling et al. 2011, Koched et al. 2013, Muhling et al. 2013). The physiological process behind this association is unknown and it has been hypothesized to be related with the ability of adults to detect salinity gradients or other processes associated with the front, which are generally related to the formation of filaments and eddies (Alvarez-Berastegui et al. 2016). The particular response observed for the number of groups with a main peak around 38.2-38.3 and another secondary peak at around 37.0 could be reflecting more a difference between the

salinities in the Balearic Islands and in the other two areas, where mean salinity is higher. However, separated modelling would be advisable to clarify this point.

Regarding the chlorophyll-a, it has been consistently observed that BFT spawning grounds are often located in areas with low CHLa concentrations (Muhling et al. 2011, Koched et al. 2013, Muhling et al. 2013, Llopiz and Hobday 2015). Our results show a peak at low chlorophyll-a values, although not for the lowest ones. Nevertheless, the extreme values typical of this variable may be obscuring somehow the results and it could be beneficial to try to refine the model response regarding this variable to see if some more clear pattern arose. Modelling the differences between areas could also add some additional light to the ecological interpretation.

Temperature is an important variable for BFT reproduction, driving the onset of the species spawning, increasing the gonadosomatic index of mature individuals, being clue for embryo and larvae development and survival and determining the spawning spatial distribution (Alemany et al., 2010; Koched et al., 2013; Muhling et al., 2013; Medina et al., 2002; Gordo and Carreras, 2014). Our results show that the temporal gradient of SST over the previous fifteen days was useful predicting the number of groups of BFT detected, in the line with what was already observed in the Balearic Islands for tuna larvae (Álvarez-Berastegui et al., 2016). A preference of BFT for the locations where the increase in temperature is around 2°C in 15 days (about 2 weeks) can be envisaged although the wide confidence interval observed for the extreme values should be further investigated to avoid any bias on the results.

The model generalizes the responses between areas. However, visual inspection of density plots of the locations where there is presence of BFT groups for the environmental variables that were selected in the final model (Figure 43) suggests that, at least for some of them, it would be advisable an in deep analysis of each area independently, in order to decipher if there could be differing ecological processes ongoing in the different areas. A range of interactions between block and environmental covariates were, however, additionally tested (results are not present in this model and available on demand) but none was retained in the model selection. A possible three-way interaction between block and year and environmental covariate can be considered in the future.

## Conclusions

The environmental information, expressed in the different covariates here explored, has good capability to provide insights of the spatio-temporal distribution of BFT in the sampled areas. The fact that these variables have more explanatory capability than location (latitude, longitude) is a sign of advancing on the right path to predict distribution of BFT and to produce relative interannual standardized indices of abundance in the future. Results also show that environmental variables, as they are provided by data producers, present caveats that have to be resolved in order to maximize modelling capabilities. One example is the capability of FSLE to discriminate segments with positive presence of BFT groups (Figure 20), but this variable presents a high percentage of NAs due to restrictions in data availability for recent years, hampering its applicability to this study. This could be resolved by establishing specific

agreements with data providers, resolving this caveat provides a promising path for a new set of environmental variables improving modelling capabilities of tuna distribution.

It is also notable that the maximum number of groups and salinity is well identified when histograms of salinity at the places where BFT were spotted are plotted (Figure 43. Density plots for locations with any spotted group of BFT by block and selected environmental variable.), showing that hydrodynamic models start to provide potential source of information reliable for modelling spatial distributions of tunas, which was not the case in previous studies (Reglero et al. 2014). Nevertheless, it is important to note that modelling approach should consider that the values at which presence is maximized may be different for the different areas. This would require a further inspection of model responses and results, and probably trying to fit a separated model in each block to avoid a possible loss of information in the response functions that can arise in the model for the three areas together (See Figure 24), otherwise mixing areas with variables that may have different response function in each area may hamper model performance and ecological interpretation.

Ecological interpretation of the response functions is also affected by the model configurations, as a priority for maximizing deviance explained by increasing the number of degrees of freedom (alternative chosen for the development of current models), may derive in complex covariate-explanatory variable relationships, as it is the case for variables such chlorophyll-a and sea surface temperature temporal gradient. In general, this study shows the complexity of developing spatial models for BFT based on environmental variability, involving aspects of biological data collection, agreements with third partners for environmental information, acquisition and prioritization on modelling objectives.

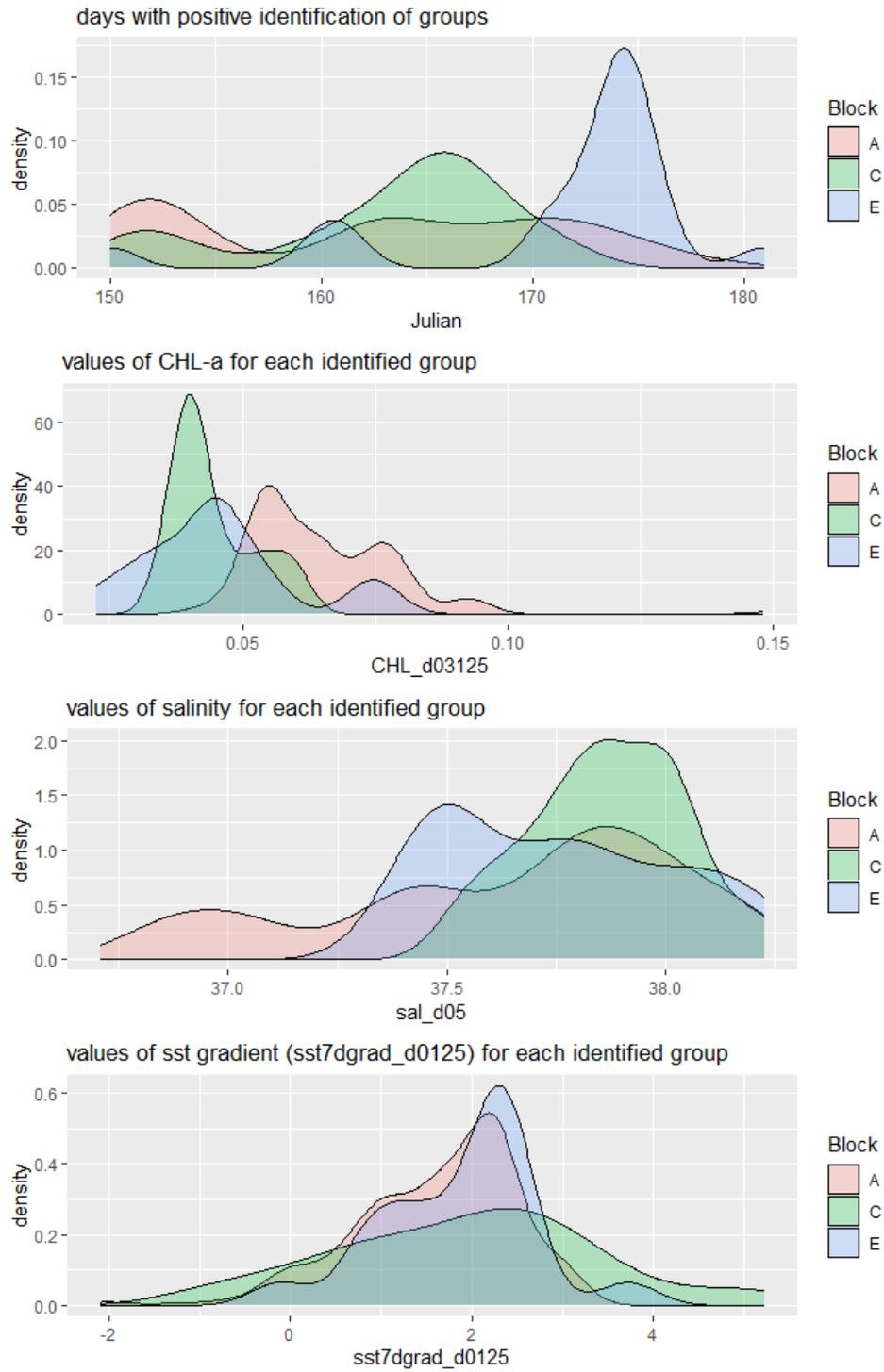


Figure 43. Density plots for locations with any spotted group of BFT by block and selected environmental variable.

## Acknowledgements

This work has been carried out under the ICCAT Atlantic-Wide Research Programme for BFT (GBYP), which is funded by the European Union, several ICCAT CPCs, the ICCAT Secretariat, and other entities (see <https://www.iccat.int/gbyp/en/overview.asp>). The content of this paper does not necessarily reflect ICCAT's point of view or that of any of the other sponsors, who carry no responsibility. In addition, it does not indicate the Commission's future policy in this area. We want to thank Dr. Hernández-Carrasco for updating the FSLE index for the aim and purposes of the present project.

## References

- Akaike, H. 1987. Factor analysis and AIC. *Psychometrika*.
- Alemaný, F., L. Quintanilla, P. Vélez-Belchí, A. García, D. Cortés, J. M. Rodríguez, M. F. de Puellas, C. González-Pola, and J. L. López-Jurado. 2010. Characterization of the spawning habitat of Atlantic bluefin tuna and related species in the Balearic Sea (western Mediterranean). *Progress in Oceanography* **86**:21-38.
- Alvarez-Berastegui, D., L. Ciannelli, A. Aparicio-Gonzalez, P. Reglero, M. Hidalgo, J. L. López-Jurado, J. Tintoré, and F. Alemany. 2014. Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution. *PLoS ONE* **9**:e109338.
- Alvarez-Berastegui, D., M. Hidalgo, M. P. Tugores, P. Reglero, A. Aparicio-González, L. Ciannelli, M. Juza, B. Mourre, A. Pascual, and J. L. López-Jurado. 2016. Pelagic seascape ecology for operational fisheries oceanography: modelling and predicting spawning distribution of Atlantic bluefin tuna in Western Mediterranean. *ICES Journal of Marine Science* **73**:1851-1862.
- Aranda, G., F. J. Abascal, J. L. Varela, and A. Medina. 2013. Spawning behaviour and post-spawning migration patterns of Atlantic bluefin tuna (*Thunnus thynnus*) ascertained from satellite archival tags. *PLoS ONE* **8**:e76445.
- Balbín, R., J. L. López-Jurado, M. d. M. Flexas, P. Reglero, P. Vélez-Velchí, C. González-Pola, J. M. Rodríguez, A. García, and F. Alemany. 2014. Interannual variability of the early summer circulation around the Balearic Islands: driving factors and potential effects on the marine ecosystem. *Journal of Marine Systems* **138**:70-81.
- Bettencourt, J. H., C. López, and E. Hernández-García. 2012. Oceanic three-dimensional Lagrangian coherent structures: A study of a mesoscale eddy in the Benguela upwelling region. *Ocean Modelling* **51**:73-83.
- Block, B. A., S. L. Teo, A. Walli, A. Boustany, M. J. Stokesbury, C. J. Farwell, K. C. Weng, H. Dewar, and T. D. Williams. 2005. Electronic tagging and population structure of Atlantic bluefin tuna. *Nature* **434**:1121-1127.
- Buckland, S. T., E. A. Rexstad, T. A. Marques, and C. S. Oedekoven. 2015. *Distance sampling: methods and applications*. Springer.
- Chudzinska, M., L. Burt, D. Borchers, D. Miller, and S. Buckland. 2021. Design-based inference to estimate density, abundance and biomass of bluefin tuna. Reanalysis of 2010-2019 Aerial Surveys. Report number CREEM-2021-03. Provided to ICCAT, September 2021 (Unpublished).
- Chudzinska, M., L. Burt, and S. Buckland. 2022. Design-based inference to estimate density, abundance and biomass of bluefin tuna. Reanalysis of 2017-2021 Aerial Surveys. Report number CREEM-2022-03. Provided to ICCAT, July 2022 (Unpublished).
- Ciannelli, L., K. Bailey, and E. M. Olsen. 2015. Evolutionary and ecological constraints of fish spawning habitats. *ICES Journal of Marine Science* **72**:285-296.

- Díaz-Barroso, L., I. Hernández-Carrasco, A. Orfila, P. Reglero, R. Balbín, M. Hidalgo, J. Tintoré, F. Alemany, and D. Álvarez-Berastegui. 2022. Singularities of surface mixing activity in the Western Mediterranean influence bluefin tuna larval habitats. *Marine Ecology Progress Series* **685**:69-84.
- Fox, J., and S. Weisberg. 2018. *An R companion to applied regression*. Sage publications.
- García, A., F. Alemany, P. Vélez-Belchí, J. L. López Jurado, D. Cortés, J. De la Serna, C. González Pola, J. Rodríguez, J. Jansá, and T. Ramírez. 2005. Characterization of the bluefin tuna spawning habitat off the Balearic Archipelago in relation to key hydrographic features and associated environmental conditions. *Col. Vol. Sci. Pap. ICCAT* **58**:535-549.
- Gordoa, A., and G. Carreras. 2014. Determination of temporal spawning patterns and hatching time in response to temperature of Atlantic bluefin tuna (*Thunnus thynnus*) in the Western Mediterranean. *PLoS ONE* **9**:e90691.
- Hazen, E. L., A. B. Carlisle, S. G. Wilson, J. E. Ganong, M. R. Castleton, R. J. Schallert, M. J. Stokesbury, S. J. Bograd, and B. A. Block. 2016. Quantifying overlap between the Deepwater Horizon oil spill and predicted bluefin tuna spawning habitat in the Gulf of Mexico. *Scientific Reports* **6**:33824.
- Hedley, S. L., and S. T. Buckland. 2004. Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics* **9**:181-199.
- Hernández-Carrasco, I., C. López, E. Hernández-García, and A. Turiel. 2011. How reliable are finite-size Lyapunov exponents for the assessment of ocean dynamics? *Ocean Modelling* **36**:208-218.
- Hijmans, R. J. 2018. raster: Geographic data analysis and modeling. R package version 2:8.
- Horvitz, D. G., and D. J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**:663-685.
- Koched, W., A. Hattour, F. Alemany, and A. García-García. 2013. Spatial distribution of tuna larvae in the Gulf of Gabes (Eastern Mediterranean) in relation with environmental parameters. *Mediterranean Marine Science* **14**:5-14.
- Laake, J., D. Borchers, L. Thomas, D. Miller, and J. Bishop. 2022. mrds: mark-recapture distance sampling. R package version 2.2.8. <https://cran.r-project.org/package=mrds>.
- Llopiz, J. K., and A. J. Hobday. 2015. A global comparative analysis of the feeding dynamics and environmental conditions of larval tunas, mackerels, and billfishes. *Deep Sea Research Part II: Topical Studies in Oceanography* **113**:113-124.
- Marques, T. A., L. Thomas, S. G. Fancy, and S. T. Buckland. 2007. Improving estimates of bird density using multiple-covariate distance sampling. *The Auk* **124**:1229-1243.
- Medina, A., F. Abascal, C. Megina, and A. Garcia. 2002. Stereological assessment of the reproductive status of female Atlantic northern bluefin tuna during migration to Mediterranean spawning grounds through the Strait of Gibraltar. *Journal of Fish Biology* **60**:203-217.
- Miller, D. L. 2022. Package 'Distance'. Versión 0.9 **6**.
- Miller, D. L., E. Rexstad, L. Thomas, L. Marshall, and J. L. Laake. 2019. Distance Sampling in R. *Journal of Statistical Software* **89**:1 - 28.
- Muhling, B. A., S.-K. Lee, J. T. Lamkin, and Y. Liu. 2011. Predicting the effects of climate change on bluefin tuna (*Thunnus thynnus*) spawning habitat in the Gulf of Mexico. *ICES Journal of Marine Science* **68**:1051-1062.
- Muhling, B. A., P. Reglero, L. Ciannelli, D. Alvarez-Berastegui, F. Alemany, J. T. Lamkin, and M. A. Roffer. 2013. Comparison between environmental characteristics of larval bluefin tuna *Thunnus thynnus* habitat in the Gulf of Mexico and western Mediterranean Sea. *Marine Ecology Progress Series* **486**:257-276.
- Pante, E., and B. Simon-Bouhet. 2013. marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* **8**.
- Pebesma, E., and R. Bivand. 2023. *Spatial data science: With applications in R*. CRC Press.

Pierce, D., and M. D. Pierce. 2019. Package 'ncdf4'. See <https://www.vps.fmvz.usp.br/CRAN/web/packages/ncdf4/ncdf4.pdf>.

R Core Team. 2022. R: A language to environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Reglero, P., L. Ciannelli, D. Álvarez-Berastegui, R. Balbín, J. L. López-Jurado, and F. Alemany. 2012. Geographically and environmentally driven spawning distributions of tuna species in the western Mediterranean Sea. *Marine Ecology Progress Series* **463**:273-284.

Reglero, P., D. P. Tittensor, D. Álvarez-Berastegui, A. Aparicio-González, and B. Worm. 2014. Worldwide distributions of tuna larvae: revisiting hypotheses on environmental requirements for spawning habitats. *Marine Ecology Progress Series* **501**:207-224.

Wood, S., and M. S. Wood. 2015. Package 'mgcv'. R package version **1**:729.

Wood, S. S. N. 2006. *Generalized Additive Models: An Introduction with R*. Chapman & Hall, UK.

Worm, B., M. Sandow, A. Oschlies, H. K. Lotze, and R. A. Myers. 2005. Global patterns of predator diversity in the open oceans. *Science* **309**:1365-1369.

## Appendix 1

GAM diagnostics are shown in Appendix Figures A1-2. Diagnostic plots for the zero-inflated model are difficult to interpret and so are not shown.

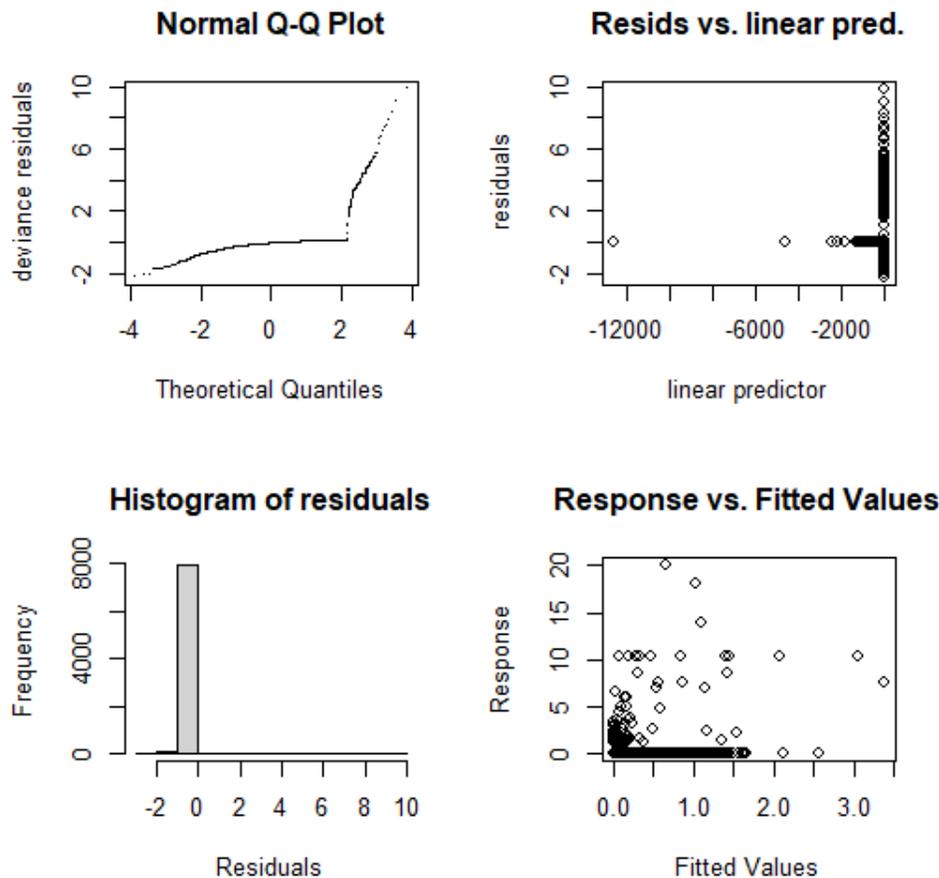


Figure A1. Diagnostic plots for the group number model

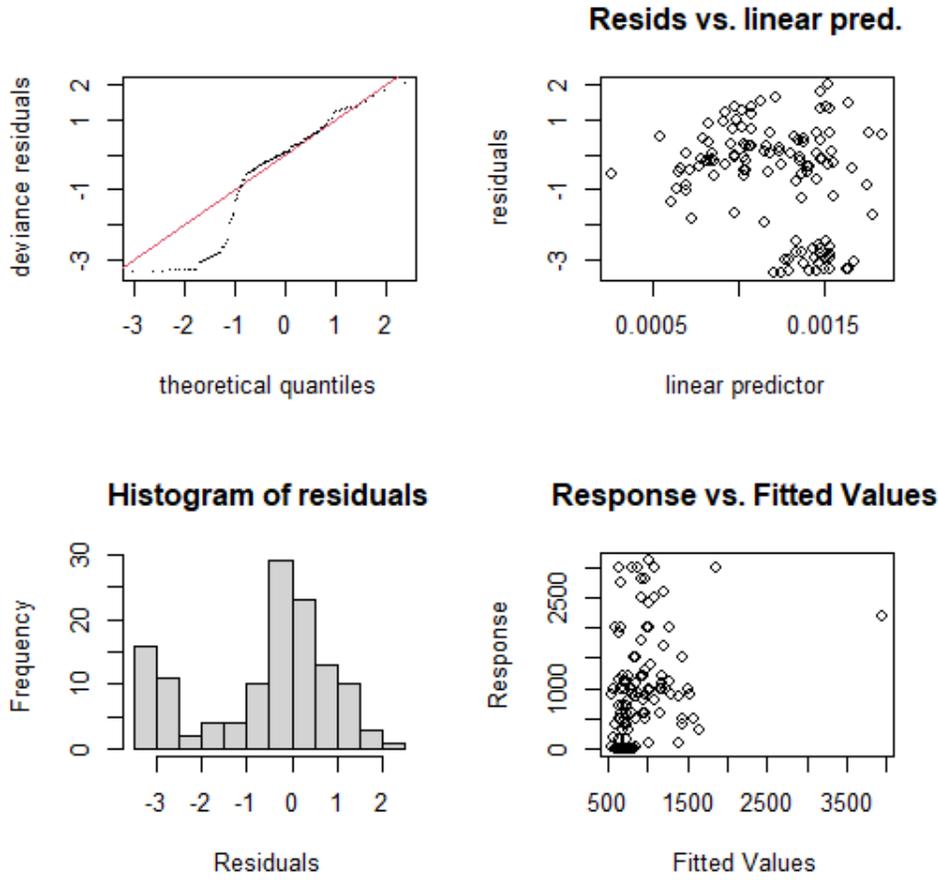


Figure A2. Diagnostic plots for the group size model.

## Appendix 2 Revised Block A design based estimates using additional observations from Group G

To allow backwards compatibility with previous estimates, additional design-based estimates were made for Block A based on an enhanced observation dataset including the large number of observations in Block G made in previous years. Assuming consistency of observation across Blocks and years this allowed a more precise estimation of the relevant detection functions.

Detection function model selection was undertaken as before.

### Results

The total number of observations were now 211 for both biomass and shoal size observations.

The best model for abundance was a half normal function on distance with company and log (size) as covariates (Figure A3).

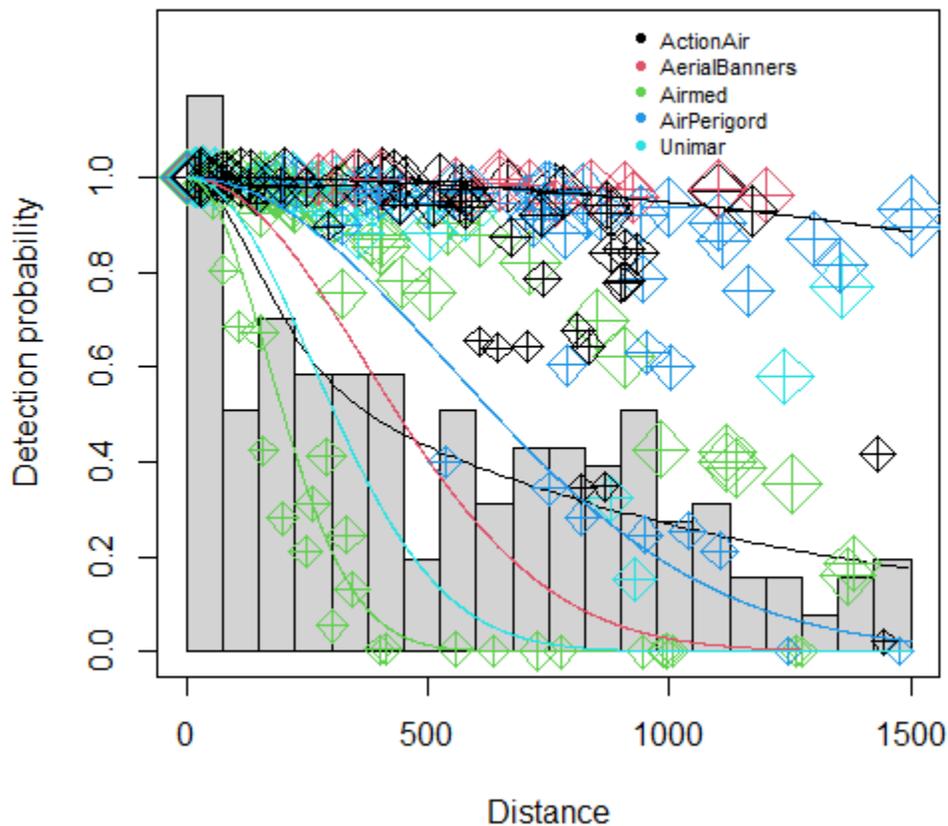


Figure A3. Histogram of observed distances, average detection function across all observations (histogram line) and detection probabilities of observed distances from best fitting model (dots) colour coded by company for abundance. Size of symbols were scaled to represent the natural log of school size.

The best model for biomass was a half normal function on distance with company and log (biomass) as covariates (Figure A3).

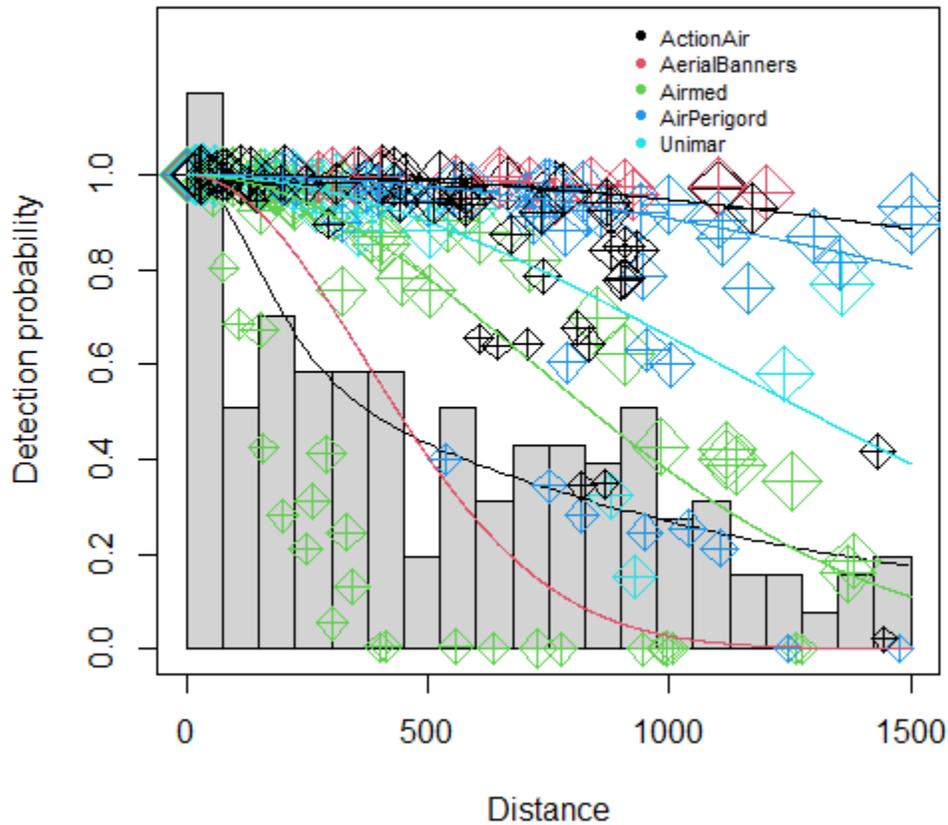


Figure A4. Histogram of observed distances, average detection function across all observations (histogram line) and detection probabilities of observed distances from best fitting model (dots) colour coded by company for abundance. Size of symbols were scaled to represent the natural log of estimated biomass.

With the detection functions estimated, abundance and biomass could be estimated as in the main text before (Table A1 and A2 respectively).

Table A2. Estimated number of individuals  $N$  (in thousands) for Block A using observations from all other blocks with 95% CIs. The same columns with 'str' apply to estimates reported in [Chudzinska et al. 2021](#), and [Chudzinska et al. 2022](#). '-' indicates that estimates previous to this study were not provided.

	Abund.	CV	LCL	UCL	Abund_str	CV_str	LCL_str	UCL_str
A-2017	55.9	0.43	24.6	127.1	49.92	0.44	21.82	114.2
A-2018	91.2	0.3	50.5	165	81.6	0.31	45.28	147.1
A-2019	86.5	0.4	40.3	185.6	75.02	0.38	36.71	153.3
A-2021	30.2	0.53	11.2	81.6	26.11	0.54	9.59	71.13
A-2022	40.4	0.38	19.3	84.6	-	-	-	-

Table A2. Estimated biomass (B, in tonnes) per block along with its standard error (SE) based on observations from all blocks, coefficient of variation (CV), lower and upper 95% confidence limits (LCL and UCL). The same columns with 'str' apply to estimates reported in [Chudzinska et al. 2021](#), and [Chudzinska et al. 2022](#). '-' indicates that estimates previous to this study were not provided.

	B	CV	LCL	UCL	B_str	CV_str	LCL_str	UCL_str
A-2017	9393	0.44	4053	21765	8001	0.45	3436	18634
A-2018	15683	0.31	8576	28680	13345	0.31	7352	24222
A-2019	13947	0.4	6467	30078	11548	0.38	5619	23734
A-2021	5701	0.53	2106	15428	4714	0.53	1750	12696
A-2022	9234	0.45	3885	21947	-	-	-	-